CERGE
Center for Economic Research and Graduate Education
Charles University Prague

Essays on Endogenous Information Acquisition in Economics

Pavel Ilinov

Dissertation

Prague, March 2023

Pavel Ilinov

Essays on Endogenous Information Acquisition in Economics

Dissertation

Prague, March 2023

Dissertation Committee

Jakub Steiner (CERGE-EI and University of Zurich; chair)

Jan Zápal (CERGE-EI)

Ole Jann (CERGE-EI)

Referees

Eugen Kováč (University of Duisburg-Essen)

Fabio Michelucci (Ca' Foscari University of Venice)

Contents

Abstract

The first chapter studies the properties of the stochastic choice data in the problems with endogenous information acquisition. We consider two types of models: (i) a rational inattention problem and (ii) a conformity game, in which fully informed players find it costly to deviate from average behavior. We show that these problems are equivalent to each other, both from the perspective of the participant and the outside observer: each individual faces identical trade-offs in both situations, and an observer would not be able to distinguish the two models from the choice data they generate. We also establish when individual behavior in the conformity game maximizes welfare.

The second chapter focuses on the flexibility of the endogenous information that an agent can acquire. We show that the principal can strictly benefit from delegating a decision to an agent whose opinion differs from that of the principal. We consider a "delegated expertise" problem in which the agent has an advantage in information acquisition relative to the principal, rather than having preexisting private information. When the principal is ex ante predisposed towards some action, it is optimal for her to hire an agent who is predisposed towards the same action, but to a lesser extent, since such an agent would acquire more information, which outweighs the bias stemming from misalignment. We show that belief misalignment between an agent and a principal is a viable instrument in delegation, performing on a par with contracting and communication in a class of problems.

The third chapter investigates the intertemporal trade-off of endogenous information in dynamic decision-making problems. We extend the classical search framework by allowing the decision-maker to choose information endogenously and flexibly. We consider a stylized problem of an interview design in which the manager chooses the best among ex ante identical candidates. The manager learns the qualities of candidates sequentially and can choose the information structure by herself at a cost. Our findings vary depending on the number of available candidates. We show that with two candidates, the optimal learning strategy is not unique: the manager may learn different information about the candidates, and there is a continuum of optimal learning strategies. However, the manager, on average, chooses candidates uniformly across all optimal strategies. On the contrary, when three candidates are available, the optimal strategy is unique. The manager engages in "cherry-picking" behavior: the difficulty of an interview decreases according to the order of candidates. Nevertheless, the manager optimally balances between the difficulty and the informativeness of an interview. In the optimum the manager

discriminates against the last candidate and chooses him least often. Additionally, we connect our findings with the serial-position effect from the psychological literature. With three alternatives the primacy effect is present: the average amount of information obtained decreases in the order of candidates. We also show that our results are robust to several extensions.

Abstrakt

V první kapitole zkoumáme vlastnosti dat stochastické volby v problémech s endogenním získáváním informací. Uvažujeme dva typy modelů: (i) model racionální nepozornosti a (ii) hru shody, kde plně informované hráče čelí nákladům při odchýlení od průměrného chování. Ukazujeme, že tyto modely jsou vzájemně ekvivalentní z pohledu účastníka i vnějšího pozorovatele: každý účastník čelí shodným kompromisům při rozhodování v obou modelech a pozorovatel není schopen rozlišit modely pouze s využitím znalosti výsledných rozhodnutí, které mohou nastat. Také zjišťujeme, kdy chování jednotlivce ve hře shody maximalizuje blahobyt.

Ve druhé kapitole se zaměřujeme na flexibilitu endogenních informací, které může agent získat. Ukazujeme, že principál může striktně mít prospěch z delegování rozhodnutí na agenta, jehož názor je odlišný od názoru principála. Zkoumáme problém „delegované expertízy", ve kterém má agent oproti principálovi výhodu při získávání informací, jako protiklad k problému s privátní informací určenou předem. Pokud má principál ex ante predispozice pro nějaké rozhodnutí, je pro něj optimální najmout agenta, který má predispozice pro stejné rozhodnutí. Predispozice agenta by neměly být ve srovnání s principálem tak silné, jelikož takový agent získá více informací, které převáží zkreslení plynoucí z neshody principála a agenta. Ukazujeme, že neshoda v názorech principála a agenta je smysluplný nástroj při delegování a srovnatelný s výsledky kontrahování a komunikování v rámci stejné třídy problémů.

Třetí kapitola zkoumá intertemporální kompromisy endogenních informací v dynamických problémech rozhodování. Rozšiřujeme klasický framework vyhledávání tím, že umožňujeme tomu, kdo rozhoduje, endogenně a flexibilně vybírat informace. Uvažujeme o stylizovaném problému designu pohovoru, ve kterém manažer vybírá toho nejlepšího z ex ante identických kandidátů. Manažer postupně zjišťuje kvalitu kandidátů a může si sám zvolit nákladní informační strukturu. Naše zjištění se liší v závislosti na počtu dostupných kandidátů. Ukazujeme, že u dvou kandidátů není optimální strategie učení jedinečná: manažer se může o kandidátech dozvědět různé informace a existuje kontinuita optimálních strategií učení. Manažer však v průměru vybírá kandidáty jednotně napříč všemi optimálními strategiemi. Naopak, když jsou k dispozici tři kandidáti, optimální strategie je jedinečná. Manažer „vybírá třešničky": obtížnost pohovoru klesá podle pořadí kandidátů. Nicméně, manažer optimálně balancuje mezi obtížností a informativností rozhovoru. V optimu manažer posledního kandidáta diskriminuje a vybírá si ho nejméně často. Naše zjištění navíc spojujeme s efektem prvenství z psychologie. U tří alternativ existuje efekt

prvenství: průměrné množství získaných informací klesá v pořadí kandidátů. Ukazujeme také, že naše výsledky jsou robustní pro několik rozšíření.

Acknowledgments

I want to thank my supervisor, Jakub Steiner, for his guidance throughout my doctoral studies. I am grateful to the other committee members, Ole Jann and Jan Zápal, for their support and critical thinking during our various discussions. I am also thankful to Fabio Michelucci, Avner Shaked, and Krešimir Žigić, who showed me the beauty of Microeconomics.

I am in debt to all CERGE-EI faculty members with whom I had a chance to interact and learn during my graduate studies. I am also extremely thankful to the members of the Academic Skills Center and especially Deborah Nováková and Andrea Downing, who were able to help me organize my writings into readable texts. I am also thankful to all CERGE-EI staff for their help with administrative duties.

I am grateful to all my coauthors, Ole Jann, Maxim Senkov, and Egor Starkov, for their patience, and I hope our collaboration will continue in the future.

I want to thank all my fellow students for sharing various ups and downs during graduate studies; it was essential. I thank my fellow student, friend, and coauthor, Andrei Matveenko, for his help when it was needed and for our endless discussions. I am thankful to all my friends outside of academia, who were always there for me. I also thank the community of my former university, EUSP. Without them, I simply would not be where I am now.

Financial support from the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement No. 770652) and support by the Charles University in Prague (GA UK projects No. 323221) is gratefully acknowledged.

All errors remaining in this text are my responsibility.

Czech Republic, Prague — Pavel

March 2023

Introduction

The unifying theme of these three chapters is analysis of problems with endogenous information acquisition. The idea that people cannot absorb all available information and pay attention only to parts of it dates back at least to Simon et al. (1971). Rather than being fully myopic, people use their ability to perceive information as scarce resource and choose second-best solutions given the degree of friction present. The theory of rational inattention, which was pioneered by Sims (2003), brought this idea to economics. The main assumption of the theory is that, before making a choice, a person optimally chooses an information structure herself, incurring greater costs for a more informative structure. In the modern world, this assumption is very natural in various settings.

In this dissertation, I use stylized models to study the implications of endogenous information acquisition on choice. In the first chapter, we analyze the structure of the stochastic choice that endogenous information acquisition leads to. The second chapter focuses on the role of the flexibility of information. The third chapter studies the intertemporal tradeoff generated by endogenous information acquisition in a search problem.

In the first paper, "An Equivalence Between Rational Inattention Problems and Complete-Information Conformity Games" (joint work with Ole Jann), which was published in *Economic Letters*, we explore the structure of the stochastic choice resulting from rationally inattentive behavior with an entropy cost function. We consider a conformity game in which a person finds it costly to deviate from the average behavior. We show that the choice data in both problems is indistinguishable from an outside observer's point of view.

Using the entropy function's proneness property, we show that strategic externalities cancel themselves out in the conformity game. Additionally, we argue that the entropy-based cost is a unique function in the additively separable class for our result to hold.

In the second paper, "Optimally Biased Expertise" (joint work with Andrei Matveenko, Maxim Senkov, and Egor Starkov), we consider a stylized model of a delegation problem. The principal delegates a decision under uncertainty to an agent who has access to information technology and is rationally inattentive. We focus on the situation in which the principal and agents to whom she can delegate a decision share the same preferences, but differ in prior beliefs. Because of the difference in prior beliefs, all agents choose different information structures. When choosing an agent, the principal must balance the amount and the direction of information an agent learns. We show that, when the principal does not internalize the cost of learning, she chooses an agent who is predisposed towards the same action, but to a lesser extent. We compare several delegating tools and show that a biased prior belief is among the most beneficial for a particular class of preferences. Our results hold irrespective of who makes a decision, the principal or an agent.

The third paper, "Is it Better to be First? Search with Endogenous Information Acquisition", extends the classical search framework by allowing the decision-maker to choose endogenous and flexible information about available alternatives. We consider a stylized model of optimal interview design, in which the manager sequentially inspects identical ex ante candidates and has to choose the best one. We study the effect of the order of inspection on choice discrimination and information acquisition. Our findings vary depending on the number of available candidates. We find no discrimination with two candidates and discrimination against the last candidate in the setup with three candidates. Depending on the number of available candidates, the manager uses information differently. With two candidates, she uses information as a perfect substitute between periods: she may obtain any non-negative amount in the first period and the rest in the second. With three candidates, the amount of information obtained decreases in the order of candidates. Moreover, the manager consumes an informative signal only in the case of low signal realizations during previous interviews. Additionally, we show that our results are robust to several extensions.

Chapter 1

An Equivalence Between Rational Inattention Problems and Complete-Information Conformity Games

Co-authored with Ole Jann (CERGE-EI).

1.1 Introduction

Models of rational inattention (RI) have become a standard tool of economic analysis. In such models, an individual chooses among a set of options, about whose properties she has imperfect information. She can acquire this information at a cost that depends on the reduction in her uncertainty, usually measured by the entropy of her beliefs. Hence, she faces a trade-off between acquiring information that allows her to make better decisions, and the cost of acquiring this information.

The RI framework is commonly understood as a model of cognitive and informational phenomena: Individuals cannot choose optimally because they lack knowledge about themselves or the world around them. The barriers they face are different, both in origin and in effect, from external or physical barriers.

We show that the standard RI model is consequentially and observationally equivalent to a complete-information game that we call "conformity game". In this game, perfectly-informed players decide between a set of options, but are punished for deviating from the average choice of all players. This cost of non-conformity is given by the Kullback-Leibler divergence between individual behavior and population average.

Our result has three economic implications. On an individual level, it shows that informational and non-informational barriers, despite originating from entirely different causes, can lead to equivalent decision problems with mathematically indistinguishable trade-offs. This opens the possibilities for further investigations into how individuals may trade off these different constraints against each other. For example, how will they decide to reduce their uncertainty if this means introducing further external constraints on their behavior?

On a population level, our result can be applied to the problem of an analyst trying to infer underlying parameters from observed choice data. We establish an observational indistinguishability: Without further evidence, the analyst would be unable to determine whether "conformist" behavior is, in fact, due to conformist pressures on the decision makers, or whether it results from a lack of information that they have to overcome by costly information acquisition. This indistinguishability holds as long as different agents in the RI model access uncorrelated sources of information.

Finally, in deriving our result, we establish a property of the class of conformity games that we consider in this paper: Individual behavior will only maximize welfare in such games if the cost of divergence is given by the Kullback-Leibler function.

A brief intuition. Consider that, on an abstract level, both models concern the behavior of individuals as part of a population. In the RI model, an agent does not know the state of the world when she makes a decision; she only knows that she lives in one of many possible states of the world, whose respective probability is given by her prior beliefs. She therefore acts as the social planner of a hypothetical population of agents corresponding to all possible states of the world.

In the conformity game, by contrast, each agent is part of an actual population of players who exert an externality on each other, since their choice determines the population average from which deviation is costly. However, as we show in proposition 1, if deviation is punished according to the Kullback-Leibler cost function, each player in the conformity game acts as if she was trying to maximize welfare. Hence, agents in both models act as social planners of a population of agents that represents all possible types.

In both cases, if costs (of learning or divergence) are high, agents will default to a kind of population average: If learning is costly, agents will mostly choose the options that are optimal under the prior, regardless of the state of the world. If divergence is costly in the conformity game, players will, in equilibrium, stick close to the average choice of other agents, regardless of what their individual preferences are.

Relationship to other research. This paper connects two strands of research. On one side, there is the literature on rational inattention, which was started by Sims (2003). Matějka and McKay (2015) describe the general solution of a static discrete-choice model with entropy cost; Caplin et al. (2022) generalize the cost functions and introduce the notion of posterior separable cost functions.

The other side consists of studies on social norms and conformity, in particular preference-based conformity (as opposed to the belief-based conformity that informational cascades induce). Bernheim and Exley (2015) discuss different mechanisms of conformity; in this paper we focus on what they describe as *preference mechanism* and not on the signaling effect of Bernheim (1994) or other inference-based theories. In our model, the norm from which it is costly to deviate is endogenous. This is a natural assumption (since norms are constructed by the society in which they exist) that has e.g. also been used by Lindbeck et al. (1999) in their study of social norms and unemployment.

In section 1.3.1, we formulate conditions under which individually optimal behavior maximizes welfare in the conformity game we analyze. Flynn and Sastry (2020) study strategic mistakes in large games and derive a similar condition for equilibrium efficiency in their setting – though without explicitly relying on the properness property of the Kullback-Leibler divergence that we use here.

1.2 Model

We first describe the basic set-up that is common to the different models we consider. We then describe the specific assumptions of each model.

1.2.1 Basic set-up

Each individual has to choose from a finite set of alternatives $\mathcal{I} = \{1, 2, \ldots, I\}$. Each individual has a type $j \in \mathcal{J} = \{1, 2, \ldots, J\}$; each type occurs with frequency μ^j where $\sum_{j=1}^{J} \mu^j = 1$. For each agent of type j, option i has payoff u_i^j. We will consider mixed as well as pure choices, so that a choice is a vector of choice frequencies $\mathbf{p}^j = (p_1^j, \ldots, p_I^j) \in$

$\Delta(\mathcal{I})$ that gives an expected payoff of

$$\pi^j(\boldsymbol{p}^j) = \sum_{i=1}^{I} p_i^j u_i^j.$$

The models that we analyze only differ in whether agents must first learn about j (and pay a cost for that) or whether they know their type j when making their choice (and then pay some cost for diverging from average choices).

1.2.2 Rational inattention problem

In this type of model, which is close to standard models of rational inattention, we assume that individuals do not know their own type j and are hence unsure about their vector of payoffs $\boldsymbol{u}^j = \left(u_1^j, \ldots, u_I^j\right)$. They can acquire information about u^j at a cost that depends linearly on the reduction in entropy of their information – this is the "mutual information" cost function, cf. Matějka and McKay (2015) and Caplin et al. (2019).

If we write θ for an agent's type, we can write $P_{j,i} = P(\theta = j | a = i)$ for the conditional probability with which an agent who chooses option i will be of type j – this value changes as an agent gathers information. Let $\boldsymbol{P}_i = (P_{1,i}, \ldots, P_{J,i})$ be the posterior belief of the agent after she has acquired information; it describes the posterior probabilities of different types given that she has found action i optimal. The cost of information is then given by the mutual information of the prior belief $\boldsymbol{\mu}$ and the vector of posterior beliefs $\boldsymbol{P}$. Since cost is given by reduction in entropy, we can express it using the Kullback-Leibler divergence[1]

$$C(\boldsymbol{P}, \boldsymbol{\mu}) = E_i \left[D_{KL}(\boldsymbol{P}_i \| \boldsymbol{\mu}) \right] = \sum_{i=1}^{I} \bar{p}_i \sum_{j=1}^{J} P_{j,i} \log\left(\frac{P_{j,i}}{\mu^j}\right)$$

where we write $\bar{p}_i = \sum_{j=1}^{J} \mu^j P_{i,j}$ for the marginal probability of action i.

The expected payoff for one agent is then given by

$$\sum_{i=1}^{I} \bar{p}_i \sum_{j=1}^{J} \left(P_{j,i} u_i^j - P_{j,i} \log\left(\frac{P_{j,i}}{\mu^j}\right)\right).$$

[1] We use the usual conventions $0 \log 0 = 0, p \log \frac{p}{0} = -\infty$ and $0 \log \frac{0}{0} = 0$.

1.2.3 Conformity game

Now consider a game that has the structure described in 1.2.1 and in which each agent perfectly knows their own payoffs, i.e., an agent of type j knows that their type is j and their payoffs are $\boldsymbol{u}^j = (u_1^j, \ldots, u_I^j)$. There is a continuum of agents and each type j continues to occur with proportion μ^j. In the absence of further constraints, the agent would simply choose the options with the highest u_i^j with probability 1. We assume, however, that the agent faces a cost of deviation from the average choice frequency $\bar{\boldsymbol{p}} = \sum_{j=1}^J \mu^j \boldsymbol{p}^j$. This cost is given by the Kullback-Leibler divergence between individual and average choice frequence, $D_{KL}(\boldsymbol{p}^j \| \bar{\boldsymbol{p}}) = \sum_{i=1}^I p_i^j \log\left(\frac{p_i^j}{\bar{p}_i}\right)$.

The agent's payoff is hence

$$\pi^j(\boldsymbol{p}^j) = \sum_{i=1}^I \left[p_i^j u_i^j - p_i^j \log\left(\frac{p_i^j}{\bar{p}_i}\right)\right]$$

We are agnostic about the origin of the deviation cost that the agent faces. It could be an intrinsic taste for conformity (as described by Bernheim and Exley (2015) as *preference mechanism*) or an external cost that is levied by some institution or society itself.

1.3 Equivalence between the models

1.3.1 Individual choice and welfare maximization in the conformity game

When making their choice $\boldsymbol{p}^j$ in the conformity game, individuals influence the average choice $\bar{\boldsymbol{p}}$ and hence exert an externality on others. We can show, however, that this externality does not distort behavior away from the welfare-maximizing set of choices, due to the properness property of the Kullback-Leibler cost function.

The statement holds for interior equilibria – each conformity game also has a (large) set of equilibria in which all players choose some options with probability zero and deviation from this probability is punished with infinite cost. Such equilibria, however, are neither welfare-maximizing nor trembling-hand stable.

Proposition 1. *Nash Equilibrium behavior maximizes welfare in the conformity game.*

Proof. Our argument is based on showing that the systems of first-order conditions in the individual problem and in the welfare maximization problem have the same set of solutions. Since the Kullback-Leibler divergence is concave in both arguments and we consider only interior equilibria, this is sufficient to show that the two optimization problems (individual and welfare) are equivalent.

Consider the Lagrangian for the problem of maximizing welfare:

$$\mathcal{L}(\boldsymbol{p}, \boldsymbol{\xi}) = \sum_{j=1}^{J} \mu^j \left(\sum_{i=1}^{I} p_i^j u_i^j - D_{KL}(\boldsymbol{p}^j \| \bar{\boldsymbol{p}}) \right) - \sum_{j=1}^{J} \xi^j (\sum_{i=1}^{I} p_i^j - 1). \tag{1.1}$$

We can rearrange the first-order condition for p_i^j as

$$\begin{aligned} &\mu^j \Big(u_i^j - \frac{\partial D_{KL}(\boldsymbol{p}^j \| \bar{\boldsymbol{p}})}{\partial p_i^j} - \sum_{k=1}^{J} \mu^k \frac{\partial D_{KL}(\boldsymbol{p}^k \| \bar{\boldsymbol{p}})}{\partial \bar{p}_i} \Big) - \xi^j = 0 \\ \Leftrightarrow \quad &\underbrace{u_i^j - \frac{\partial D_{KL}(\boldsymbol{p}^j \| \bar{\boldsymbol{p}})}{\partial p_i^j}}_{\text{primary effect}} = \frac{\xi^j}{\mu^j} + \underbrace{\sum_{k=1}^{J} \mu^k \frac{\partial D_{KL}(\boldsymbol{p}^k \| \bar{\boldsymbol{p}})}{\partial \bar{p}_i}}_{\text{externality effect}} \end{aligned} \tag{1.2}$$

The primary effect is the effect of an agent's action on their own payoff. In the individual problem of the agent j, it equals the Lagrange multiplier. Hence, if we can show that the right-hand side of equation (1.2) depends only on j, then we can simply rescale the Lagrange multiplier to transform one system of equations into the other. For this we need to show that the externality effect is invariant in i.

Consider, for a moment, an auxiliary optimization problem in which the social planner takes all $\boldsymbol{p}^j$ as given, but can replace $\bar{p}$ in the cost function with any other vector in order to minimize total cost. This is equivalent to finding the $\boldsymbol{q}$ that solves the problem

$$\arg\min_{\boldsymbol{q} \in \Delta(\mathcal{I})} \sum_{j=1}^{J} \mu^j \sum_{i=1}^{I} p_i^j \log \left(\frac{p_i^j}{q_i} \right) \tag{1.3}$$

or

$$\arg\min_{\boldsymbol{q} \in \Delta(\mathcal{I})} \sum_{i=1}^{I} \bar{p}_i \log q_i.$$

This expression is minimized exactly by $\boldsymbol{q} = \bar{\boldsymbol{p}}$ due to the properness property of entropy.[2] Now consider the externality effect in expression (1.2), which is the effect of

[2]Proper scoring rules have been used in economics for belief elicitation. cf. Savage (1971) for the classical reference, and Tsakas (2020) for the recent application to the decision problems with incomplete information. For economic applications of the properness property of entropy, cf. the proof of lemma 2 in Steiner et al. (2017).

$\bar{p}_i$ on the total cost at the point $\bar{\boldsymbol{p}}$. But if $\bar{\boldsymbol{p}}$ is already optimal, that must mean all $\bar{p}_i$ have the same marginal effect, and hence the externality effect is the same for all i. This means we can transform the individual maximization problem into the welfare maximization problem and vice versa, and completes the proof. □

Note that for $I > 2$, it follows from the result of Aczél and Pfanzagl (1967) that the Kullback-Leibler divergence is the only additively separable cost function for which this proposition is true.

1.3.2 Equivalence

We are now ready to show our main result: the equivalence between rational inattention problems, as described in section 1.2.2 and complete-information conformity games, as described in section 1.2.3.

Proposition 2 (Equivalence)**.** *Consider a rational inattention problem with a finite number of options and entropy cost of information. There then exists a conformity game in which the deviation cost is given by the Kullback-Leibler divergence and which has equivalent payoffs, trade-offs and equilibrium behavior, and vice versa: for any such given conformity game there exists a related rational inattention problem with a finite number of options and an entropy cost of information.*

Proof. We will show that the individual maximization problem of the RI model and the welfare maximization problem of the conformity game can be transformed into each other. The result then follows in combination with proposition 1.

We exploit that $\frac{P_{j,i}}{\mu^j} = \frac{p_i^j}{\bar{p}_i}$ (Bayes' rule), the fact that $P(a = i) = \bar{p}_i$ (by definition) and the martingale property of expectations to transform:

$$\begin{aligned}\sum_{i=1}^{I} \bar{p}_i \sum_{j=1}^{J} \left[P_{j,i} u_i^j - P_{j,i} \log\left(\frac{P_{j,i}}{\mu^j}\right)\right] &= \sum_{i=1}^{I} \bar{p}_i \sum_{j=1}^{J} \left[\frac{p_i^j \mu^j}{\bar{p}_i} u_i^j - \frac{p_i^j \mu^j}{\bar{p}_i} \log\left(\frac{p_i^j}{\bar{p}_i}\right)\right] \\ &= \sum_{j=1}^{J} \mu^j \sum_{i=1}^{I} \left[p_i^j u_i^j - p_i^j \log\left(\frac{p_i^j}{\bar{p}_i}\right)\right]\end{aligned}$$

where the first expression is the individual maximization problem in the RI problem and the last is welfare in the conformity game. □

Formally, we exploit that the Kullback-Leibler divergence is a function of the ratio of two probabilities. This allows us to use Bayes' rule to transform conditional probabilities of one type (how likely a state of the world is for a given action) into conditional probabilities of another type (how likely an action is for a given type).

Note that this technique is not limited to Kullback-Leibler cost functions. It would apply for any f-divergence cost function (Rényi, 1961). However, for other f-divergence cost functions, proposition 1 does not apply; hence there is only an equivalence between an individual RI problem and the social planner problem of a conformity game, but not between the individual problems of the two models.

1.A The Necessity of Kullback-Leiber Costs

We will sketch a proof showing that our result in proposition 1 is limited to conformity games with Kullback-Leibler cost functions (among the set of additively separable cost functions).

Proposition 1.1. *For $I \geq 3$, the result from proposition 1 is not true for any other additively separable cost function than the Kullback-Leibler divergence (and its linear transformations).*

Proof. In the proof of proposition 1, we have exploited that $\bar{p}$ is the minimizer of total cost. Now assume that instead of the Kullback-Leibler function, we were to work with a general additively separable cost function $C(\boldsymbol{p}^j, \boldsymbol{\mu}) = \sum_{i=1}^{I} p_i^j (f(p_i^j) - f(q_i))$ where f f is a continuously differentiable function. Then expression (1.3) becomes

$$\begin{aligned}\underset{\boldsymbol{q}\in\Delta(\mathcal{I})}{\arg\min}\Big\{\sum_{j=1}^{J}\mu^j\sum_{i=1}^{I}p_i^j(f(p_i^j)-f(q_i))\Big\} &= \underset{\boldsymbol{q}\in\Delta(\mathcal{I})}{\arg\max}\Big\{\sum_{j=1}^{J}\mu^j\sum_{i=1}^{I}p_i^j f(q_i)\Big\}\\ &= \underset{\boldsymbol{q}\in\Delta(\mathcal{I})}{\arg\max}\sum_{i=1}^{I}\bar{p}_i f(q_i).\end{aligned}$$

(Aczél and Pfanzagl, 1967) show that for $I \geq 3$, this has the solution $q_i = \bar{p}_i$ (under the constraint that $\sum_{i=1}^{I} q_i = 1$) if and only if $f(q_i) = c_1 + c_2 \log q_i$. Hence, if the cost function C were to take a form that is not a linear transformation of D_{KL}, then p_i^j would have an indirect as well as a direct effect on the welfare maximization problem (1.1), and the welfare-maximizing choice profile of the conformity game would not be a Nash Equilibrium.

$\square$

Chapter 2

Optimally Biased Expertise

Co-authored with Andrei Matveenko (University of Mannheim, Department of Economics), Maxim Senkov (CERGE-EI), Egor Starkov (University of Copenhagen, Department of Economics)

2.1 Introduction

Delegation is a valuable management tool widely used in business structures, political organizations, and other economic contexts. Firm owners delegate operational decisions to managers, politicians delegate to their advisors, grant funders delegate award decisions to experts in the field, and people rely on advice from financial and tax advisors. It is also true that in many of these scenarios, the experts do not have much preexisting knowledge about the case they consider, but rather use their expertise to more easily acquire additional information to make the best decision.[1]

What happens if the preferences or the beliefs about the fundamentals are not fully aligned between the principal and the agent? The common wisdom (see, e.g., Holmström [1980]) suggests that this leads to a conflict, since from the principal's point of view, the agent then makes suboptimal decisions.[2] However, in this paper we show that when

[1]E.g., Graham et al. (2015) show that delegation tends to be used when the decision-making demands more evidence that the delegatee can provide. Alternatively, the choice to delegate a decision is often associated with a volatile environment that a delegator faces (Foss and Laursen, 2005; Ekinci and Theodoropoulos, 2021), so any knowledge quickly becomes obsolete.

[2]This is supported by the empirical evidence: e.g., Hoffman et al. (2018) find that inefficiencies in HR managers' hiring decisions can be a result of their biased preferences. Kennedy (2016) presents evidence that the principals take the conflict of interest into account when selecting the expert.

the agent has to actively acquire information, as opposed to having a preexisting informational advantage, there is more to this story. In particular, such a misalignment can then benefit the principal by encouraging the agent to acquire more information than an aligned agent would.

We consider a delegation model in which the payoffs from different actions depend on the unknown state of the world, and the principal (she) and the agent (he) have misaligned beliefs about the state of the world. The agent initially has no private information about the state, but can acquire costly information about it, which would improve the quality of the decision made (this setting has been labeled by Demski and Sappington [1987] as "delegated expertise"). The agent's cost of learning is not internalized by the principal, and her own cost of learning is prohibitively high.

We show that the principal benefits the most from delegating to an agent who ex ante is more uncertain than she is about what the best course of action is (shown in different contexts by Propositions 2.1, 2.2, and 2.3). This is because the more uncertain the agent is, the more he learns about the state, and the better his action fits the state – which benefits the principal. This, however, has to be balanced against the channel described above: any kind of misalignment between the principal and the agent leads to a bias in the agent's decisions, compared to what the principal would prefer. Therefore, the principal ends up hiring an agent who is more uncertain than she is and thus conducts a more thorough investigation than an aligned agent would, – but who still shares her action predispositions to some extent (i.e., favors the same action ex ante). This result holds regardless of who has the final decision rights: the optimal delegation strategy is the same whether the principal delegates the decision rights to the agent or merely expects a recommendation on the optimal course of action (Proposition 2.10).

The presence of the principal's trade-off between the amount of information acquired by an agent and the bias in his resulting decisions (Section 2.4) relies upon the flexible information acquisition technology. We use the Shannon model of discrete rational inattention (see, e.g., Matějka and McKay [2015], and Caplin et al. [2019]) to provide this dimension of richness.[3] According to the Shannon model, an agent can choose any signal structure but has to pay a cost proportional to the expected entropy reduction.[4] The

[3]More precisely, in the analysis, we use the model of finding the best alternative, studied in Caplin et al. [2019]. We consider the Shannon model with general preferences in Section 2.7.

[4]The entropy parametrization leads to the information cost being dependent on the prior belief, even keeping the signal structure constant. This has been one of the critiques of the Shannon model (see Mensch [2018]). Such a cost function has, however, been rationalized in both information theory as a cost function arising from the optimal encoding problem (see Cover and Thomas [2012]), and decision

choice of the signal in this model depends on the agent's prior belief: an agent whose prior is skewed towards some state of the world chooses a signal which is relatively more informative regarding that state and thus allows him to make a better decision in that state. This dimension of flexibility is what leads to bias in the final decisions when the agent's prior belief is not aligned with the principal's.

The main conclusion of our paper is that delegation to an agent with misaligned beliefs is an instrument that is available – and valuable – to the principal. Not only that, but in our setting this instrument can perform as well as action-contingent payments, while bearing no cost for the principal (Proposition 2.5), and cannot be improved on by using outcome-contingent payments (Proposition 2.7). Further, this instrument is typically better than restricting the agent's choice set (Proposition 2.8). This benefit of misalignment challenges the opinion that disagreement between the principal and the agent inevitably leads to a conflict, and thus the principal should seek to hire an agent who is most aligned with her preferences and beliefs (see Holmström [1980]; Crawford and Sobel [1982]; Prendergast [1993]; Alonso and Matouschek [2008]; Egorov and Sonin [2011]; Che et al. [2013] for some examples of such a message).

We use a heterogeneous priors approach to model the misalignment between the principal and the agents. Thus, we differ from the common prior assumption, and our approach violates the so-called "Harsanyi doctrine" (Morris (1995)). We allow our agents to have heterogeneous prior beliefs and, therefore, "agree to disagree". In our model, prior beliefs are observable, verifiable[5], and heterogeneity is common knowledge. We interpret such heterogeneity as our agents having different opinions or views about the unknown payoff relevant state. Heterogeneity in prior beliefs cannot be fully explained by the different information that agents may have. The prior beliefs in our model are publicly observable, and there are no signaling concerns regarding them. The agents do not want to revise their prior beliefs if there is a difference: the disagreement persists.

We take two applied interpretations of heterogeneity in our model. The first relates to the bilateral relationship. For example, when an authority in the public organization wants to find the best expert for delegating a decision. Our findings provide conditions for the set of experts when delegation is beneficial and also an upper bound for expected

theory as arising naturally from Wald's sequential sampling model (see Hébert and Woodford [2019]). In turn, the Shannon model has been shown to work as a microfoundation of the logit choice rule commonly used in choice estimation ((Matějka and McKay, 2015)).

[5]Verifiability can come, for example, from reputational concerns. In the analysis, we abstract from the strategic incentives that agents can have by declaring their prior beliefs and focus on the delegation mechanism itself.

gains in such a delegation. Moreover, suppose it is relatively straightforward for the authority to order available experts in their opinions. In that case, we provide a useful *directional* behavioral tool: the authority should look for an expert who shares similar views but is more uncertain. See, e.g., Li and Suen (2004) for a similar application of heterogeneous beliefs in the delegation model.

Our second interpretation covers large organizations. We take heterogeneous priors as the different views of the people in organizations, e.g., research teams, firms, and political parties. Our results support a diversity of views in such organizations.[6] We show that the leader of an organization can benefit from having workers with different views. We characterize the useful diversity strategy for the leader: she benefits from having workers with slightly more *moderate* views. Our problem is static and one-shot. Although the optimal agent is unique in our model, for other decision problems, the leader may have different opinions and, therefore, benefits from having workers with other views in the organization. Banerjee and Somanathan (2001) also study the impact of the heterogeneous beliefs in large organizations in the information aggregation model.

Although in our paper we focus on the misalignments in beliefs, another potential source of heterogeneity is a difference in preferences. We show that the principal can equivalently use the misalignment in preferences rather than the misalignment in beliefs (see Theorem 2.1). Namely, Proposition 2.5 states that the best delegation outcome can be implemented by hiring an agent with either optimally misaligned beliefs, or optimally misaligned preferences (or, equivalently, offering action-contingent compensation). This result has a mirror implication for the empirical literature estimating discrete choice models: Theorem 2.1 implies that the observed choice probabilities alone do not allow an external observer to jointly identify the decision maker's beliefs and preferences in our setting.

The main conclusion of our paper is that delegation to an agent with misaligned beliefs is an instrument that is available – and valuable – to the principal. Not only that, but in our setting this instrument can perform as well as action-contingent payments, while bearing no cost for the principal (Proposition 2.5), and cannot be improved on by using outcome-contingent payments (Proposition 2.7). Further, this instrument is typically better than restricting the agent's choice set (Proposition 2.8). This benefit of misalignment challenges the opinion that disagreement between the principal and the agent inevitably leads to a conflict, and thus the principal should seek to hire an agent

[6] For the recent review on diversity in organizations, see Shore et al. (2009)

who is most aligned with her preferences and beliefs (see Holmström [1980]; Crawford and Sobel [1982]; Prendergast [1993]; Alonso and Matouschek [2008]; Egorov and Sonin [2011]; Che et al. [2013] for some examples of such a message).

Our paper mainly connects to the literature on delegation, mainly to the problems of "delegated expertise", in which delegation takes place not to an agent with some pre-existing private information, but rather to an agent whose goal is to *acquire* relevant information. The assumption is that the agent's expertise allows him to gather information at a lower cost than what the principal would have to incur. The seminal paper in the field is by Demski and Sappington (1987), who explore a contracting problem in a setting in which the agent chooses between a finite number of signal structures. Lindbeck and Weibull (2020) extend this analysis to a rationally inattentive agent (who can acquire any information, subject to entropy costs). Szalay (2005) shows that restricting the agent's action set could be a useful tool in such a setting, since banning an ex ante optimal "safe" action can nudge the agent to acquire more information about which of the risky actions is the best. Our grand message is similar: the principal is willing to sacrifice something in exchange for the agent acquiring more information, but we present a different channel through which the principal can achieve this.

The closest to our paper is contemporary work by Ball and Gao (2021). They consider a model of delegated expertise and demonstrate a result similar to that of Szalay (2005): that banning the ex ante safe actions can lead to more information acquisition by the agent, which benefits the principal. However, where Szalay (2005) looks at the scenario in which the principal's and the agent's preferences coincide ex post (i.e., net of information costs), Ball and Gao (2021) explore a model with misaligned preferences and show that the principal may benefit from some misalignment between her preferences and those of the agent. In their setting, this is due to divergence between the principal's and the agent's ex ante optimal actions (due to preference misalignment), which makes banning the ex ante agent-preferred action less costly for the principal. Our paper suggests a different channel through which misalignment may incentivize the agent's information acquisition: we show, using a flexible information acquisition framework, that misalignment can lead to more information acquisition by the mere virtue of the agent being more uncertain about what the optimal action is.

The misalignment in prior beliefs is also studied by Che and Kartik (2009). They analyze a delegated expertise game in which the principal retains the decision rights: i.e., after acquiring the relevant information, the agent must communicate it to the principal, who then makes a decision she believes is optimal. They show that the need to communicate may also incentivize an agent with misaligned preferences to acquire more information, in order to more effectively persuade the principal about which action needs to be taken. Their conclusion does, however, rely on an inflexible information acquisition structure (the agent's probability of observing the true state of the world is increasing in his effort). It appears likely that in a more flexible model, a misaligned agent would acquire not *more*, but rather *different* information in order to be persuasive.

The remainder of the paper is organized as follows: we present a simple example that demonstrates the main effect in Section 2.2. Section 2.3 formulates the main model, which is then analyzed in Section 2.4 for the special case of binary states and actions, while Section 2.5 analyzes the general problem. Section 2.6 compares misaligned beliefs as a delegation tool to other tools, such as misaligned preferences, payments, and restricting the action set. Section 2.7 explores a number of extensions of the baseline model, and Section 3.6 concludes.

2.2 Illustrative example

This section presents a simple example with inflexible learning and demonstrates how a misalignment in beliefs between a principal and an agent may benefit the principal. The full model is introduced in Section 2.3.

Consider a president (a principal, she) who needs to appoint a minister or a head of a government agency (an agent/expert, he) to solve a particular policy issue, e.g., "the green transition" policy, or the antitrust policy regarding large tech companies. Suppose that there are two policies to choose from: $a \in \{L, R\}$.[7] Which policy is optimal depends on the state of the world $\omega \in \{l, r\}$, which is initially unknown to both the principal and the agent(s). Suppose further that the principal and the agent have a common interest in implementing the correct policy.[8] In particular, the utility $u(a, \omega)$ all players derive from policy a in state ω is given by $u(L, l) = u(R, r) = 1$ and $u(L, r) = u(R, l) = 0$. We denote the probability that the principal's prior belief assigns to state $\omega = r$ as $\mu_p \in [0, 1]$. The agent can choose whether to learn the state at cost $c < 1/2$ or not, and we assume that this cost is neither internalized, nor compensated for by the principal.

[7] This binary model is common in the delegation literature, see e.g. Li and Suen (2004) with a slightly different informal story.

[8] In this example we focus on misalignment in beliefs; see Section 2.6.1 for the discussion of misalignment in preferences.

The principal's problem is to select the best agent to delegate the decision to. There are many experts available to the principal, who differ in their prior beliefs about the state of the world: $\mu \in [0, 1]$. For example, she can delegate a decision to an expert who is certain that $\omega = r$ (i.e., $\mu = 1$), or to an agent who has the same opinion as her ($\mu = \mu_p$), or to the most uncertain agent ($\mu = 1/2$). The expert's prior belief is observable by the principal, which could be due to him having an established reputation for having a certain position on the question at hand. We assume that the principal cannot use monetary transfers and/or restrict the set of actions available to the agent (see Section 2.3 for a discussion of this assumption).

The timing is as follows: the principal chooses an agent based on his prior belief μ from a pool of agents $\mathcal{M} = [0, 1]$; then the agent chooses whether to learn the state ω at cost c, and subsequently implements policy a^* preferred given his prior belief and the acquired signal.

The agent's expected utility when he learns the state and when he does not is given by, respectively:

$$\mathbb{E}[u(a^*, \omega)|\mu] - c = 1 - c,$$
$$\mathbb{E}[u(a^*, \omega)|\mu] = \max\{\mu, 1 - \mu\}.$$

Hence the willingness to pay for information of an agent with prior belief μ is given by

$$\underbrace{1 - c}_{\text{with info}} - \underbrace{\max\{\mu, 1 - \mu\}.}_{\text{without info}}$$

Since this amount is decreasing in μ when $\mu \geq 1/2$ and increasing when $\mu \leq 1/2$, it is maximized at $\mu = 1/2$. Namely, the most uncertain agent is willing to pay the most for information: if an agent with some prior μ chooses to obtain information, then an agent with prior $\mu = 1/2$ would also choose to learn the state. Therefore, the principal (weakly) prefers to delegate to the most uncertain agent, since such an agent acquires the most information. The next simple proposition formalizes the result, with the argument above serving as its proof.

Proposition 2.1. *In the equilibrium of the described model, it is (weakly) optimal for the principal to delegate the decision to the most uncertain agent: $\mu^* = 1/2$.*

More specifically, the principal is indifferent between all agents with $\mu \in [c, 1-c]$, since all such agents learn the state, and strictly prefers them to any agent with $\mu \notin [c, 1-c]$, since the latter simply take their ex ante preferred action without learning the state. Hence if the principal's prior belief is $\mu_p < c$ or $\mu_p > 1-c$, it is strictly worse for the principal to hire an aligned agent than to hire a more uncertain agent with $\mu \in [c, 1-c]$.

The restricted setting of this example, however, cannot illustrate the bias introduced in the agent's decisions by the misalignment. In the next section, we proceed to the full model, which demonstrates that there also exists a countervailing preference for a less misaligned agent due to the bias in actions that this misalignment introduces. Such a preference arises due to the agent being able to acquire information flexibly.

2.3 Model

2.3.1 Concepts and Definitions

Consider a principal (she) who would like to implement an optimal decision that depends on the unknown state of the world. To choose the best course of action, the principal delegates the decision to an expert (an agent, he), who can acquire information about what the optimal decision is.[9] There are many experts available to the principal, and all experts have a common interest with the principal, but differ in their opinions on the issue.[10] Experts with different initial opinions would acquire different information, and thus possibly make different final decisions. The principal is thus concerned with finding the best agent for the job.[11]

The above can be modeled as a game played between a principal and a population of agents. In particular, let $\mathcal{A}$ denote the set of actions with a typical element a, and Ω denote the set of states with a typical element ω. The principal has a prior belief $\mu_p \in \Delta(\Omega)$, where $\Delta(\Omega)$ denotes the set of all probability distributions on Ω. Every agent in the population has some prior belief $\mu \in \Delta(\Omega)$, which is observable and verifiable, e.g., due to the reputation concerns (i.e., agents needing to publicly establish a particular stance on a broad policy question for sake of earning, and subsequently capitalizing on,

[9] An alternative would be to ask the agent to learn about the state and report the findings to the principal, who then makes the decision. This version is explored in Section 2.7.2, which demonstrates that communication is equivalent to delegation in our setting (barring the equilibrium multiplicity).

[10] In the "green transition" policy example, the experts would differ in their stance on the severity of the climate threat.

[11] Our results can, alternatively, be interpreted as comparative statics for a game between a principal and a given agent with some fixed misalignment, w.r.t. the degree of misalignment.

a specific reputation). In what follows, we refer to an agent according to his prior belief. Let $\mathcal{M} \subseteq \Delta(\Omega)$ denote the set of prior beliefs of all agents in the population.[12,]

The terminal payoff that both the principal and the agent selected by the principal receive when action a is chosen in the state ω is given by $u(a, \omega)$. Prior to making the decision, the selected agent can acquire additional information about the realized state. We assume that the agent can choose any signal structure defined by the respective conditional probability system $\phi : \Omega \to \Delta(\mathcal{S})$, which prescribes a distribution over signals $s \in \mathcal{S}$ for all states $\omega \in \Omega$, where $\mathcal{S}$ is arbitrarily rich. The information is costly: when choosing a signal structure ϕ, the agent must incur cost $c(\phi, \mu)$ that may depend on the informativeness of the signal and the agent's prior belief μ.[13]

The cost function we consider is the Shannon entropy cost function used in rational inattention models (Matějka and McKay, 2015). In this specification, the cost is proportional to the expected reduction in the entropy of the agent's belief resulting from receiving the signal. Namely, let $\eta : \mathcal{S} \to \Delta(\Omega)$ denote the agent's posterior belief system, obtained from μ and ϕ using the Bayes' rule. The cost can be defined as

$$\begin{aligned} c(\phi, \mu) \equiv \lambda \Bigg(& -\sum_{\omega \in \Omega} \mu(\omega) \ln \mu(\omega) + \\ & + \sum_{\omega \in \Omega} \sum_{s \in \mathcal{S}} \left(\sum_{\omega' \in \Omega} \mu(\omega') \phi(s|\omega') \right) \eta(\omega|s) \ln \eta(\omega|s) \Bigg), \end{aligned} \tag{2.1}$$

where $\lambda \in \mathbb{R}_{++}$ is a cost parameter.[14] We assume that the principal does not internalize the cost of learning, and the agent fully bears this cost. The main interpretation (shared by, e.g., Lipnowski et al. [2020]) of this assumption is that the cost reflects the cognitive process of the agent. Information acquisition costs thus lead to moral hazard, with the agent potentially not willing to acquire the amount of information desired by the principal. This is the main conflict between the two parties in our model.

In line with the delegation literature, we assume that the principal cannot use monetary or other kinds of transfers to manage the agent's incentives. This is primarily because learning is non-contractible in most settings – indeed, it is difficult to think of a

[12]For most of the results we assume that the population of agents is rich enough to represent the whole spectrum of viewpoints: $\mathcal{M} = \Delta(\Omega)$.

[13]Similar to, e.g., Alonso and Câmara (2016), we assume that the agent and the principal share the understanding of the signal structure. Combined with them having different (subjective) prior beliefs over states, this implies they would also have different (subjective) posterior beliefs if both observed the signal realization.

[14]We also follow the standard convention and let $0 \ln 0 = 0$.

setting, in which a learning-based contract could be enforceable, i.e., either the principal or the agent could demonstrate beyond reasonable doubt exactly how much effort the agent has put into learning the relevant information, and what kind of conclusions he has arrived at. A simpler justification of the no-transfer assumption could be that such transfers are institutionally prohibited in some settings.[15] We do allow for some classes of transfers in Section 2.7 and show that even in those settings where contracting is feasible, it does not necessarily perform better than hiring an agent with a misaligned belief.

The game proceeds as follows. In the first stage, the principal selects an agent from the population based on the agent's prior belief μ. In the second stage, the selected agent chooses signal structure ϕ and pays cost $c(\phi, \mu)$. In the third stage, the agent receives signal s according to the chosen ϕ and selects action a given s. Payoffs $u(a, \omega)$ are then realized for the principal and the agent.

The following subsections describe the respective optimization problems faced by the principal and her selected agent, and define the equilibrium concept.

2.3.2 The Agent's Problem

The agent selected by the principal chooses a signal structure $\phi : \Omega \to \mathit{\Delta}(\mathcal{S})$ and a choice rule $\sigma : \mathcal{S} \to \mathcal{A}$ to maximize his expected payoff net of the information costs. The agent's objective function is

$$\mathbb{E}[u(a, \omega)|\mu] - c(\phi, \mu) = \sum_{\omega \in \Omega} \mu(\omega) \sum_{s \in \mathcal{S}} \phi(s|\omega) u(\sigma(s), \omega) - c(\phi, \mu),$$

The agent's problem can then be written down as

$$\max_{\phi, \sigma} \left\{ \sum_{\omega \in \Omega} \mu(\omega) \sum_{s \in \mathcal{S}} \phi(s|\omega) u(\sigma(s), \omega) - c(\phi, \mu) \right\}. \tag{2.2}$$

Lemma 1 in Matějka and McKay (2015) shows that problem (2.2) with entropy cost function can be reframed as a problem of selecting a collection of conditional choice probabilities. This reformulation is presented in Section 2.3.5.

[15]See Laffont and Triole (1990); Armstrong and Sappington (2007); Alonso and Matouschek (2008) for some examples and discussion of such settings.

2.3.3 The Principal's Problem

The principal's problem is to choose an agent based on his prior belief $\mu \in \mathcal{M}$ in order to maximize her expected utility from the action eventually chosen by the agent (where $\mathbb{E}_p$ denotes the expectation w.r.t. the principal's belief μ_p). Her objective function is

$$\mathbb{E}[u(a,\omega)|\mu_p] = \sum_{\omega\in\Omega} \mu_p(\omega) \sum_{s\in\mathcal{S}} \phi(s|\omega)u(\sigma(s),\omega),$$

so her optimization problem can be written down as

$$\begin{aligned} &\max_{\mu} \left\{ \sum_{\omega\in\Omega} \mu_p(\omega) \sum_{s\in\mathcal{S}} \phi_\mu(s|\omega)u(\sigma_\mu(s),\omega) \right\}, \\ &\text{s.t. } (\phi_\mu, \sigma_\mu) \text{ solves (2.2) given } \mu. \end{aligned} \tag{2.3}$$

where the choice of agent μ affects the signal structure ϕ_μ and the choice rule σ_μ chosen by the agent. Therefore, the principal's problem is effectively that of choosing a pair (ϕ, σ) from a menu given by the agents' equilibrium strategies.

2.3.4 Equilibrium Definition

We now present the equilibrium notion used throughout the paper; the discussion follows.

Definition 2.1 (Equilibrium)**.** An equilibrium of the game is given by $(\mu^*, \{\phi^*_\mu, \sigma^*_\mu\}_{\mu\in\mathcal{M}})$: the principal's choice $\mu^* \in \mathcal{M}$ of the agent who the task is delegated to and a collection of the agents' information acquisition strategies $\phi^*_\mu : \Omega \to \mathit{\Delta}(\mathcal{S})$ and choice rules $\sigma^*_\mu : \mathcal{S} \to \mathcal{A}$ for all $\mu \in \mathcal{M}$, such that:

1. ϕ^*_μ and σ^*_μ constitute a solution to (2.2) for every $\mu \in \mathcal{M}$;

2. μ^* is a solution to (2.3) given $(\phi^*_\mu, \sigma^*_\mu)$.

Note that the above effectively defines a Subgame-Perfect Nash Equilibrium. While our game features incomplete information (about the state of the world chosen by Nature), and the players' beliefs play a central role in the analysis, problem formulations (2.2) and (2.3) allow us to treat these beliefs as just some exogenous functions entering the terminal payoff functions. This is primarily because one player's actions do not affect another player's beliefs in this game, hence a belief consistency requirement is not needed (however, we do require internal consistency in that the agent's posterior belief η is obtained by updating his prior belief μ via Bayes' rule given his requested signal structure ϕ).

2.3.5 Preliminary Analysis

Matějka and McKay (2015) show that with entropy costs, the agent's problem of choosing the information structure and choice rule can be reduced to the problem of choosing the conditional action probabilities. Namely, the maximization problem of the agent can be rewritten as that of choosing a decision rule $\pi : \Omega \to \Delta(\mathcal{A})$ (which is a single state-contingent action distribution, as opposed to the combination of a signal strategy $\phi : \Omega \to \Delta(\mathcal{S})$ and a choice rule $\sigma : \mathcal{S} \to \mathcal{A}$):

$$\max_{\pi} \left\{ \sum_{\omega \in \Omega} \mu(\omega) \left(\sum_{a \in \mathcal{A}} \pi(a|\omega) u(a, \omega) \right) - c(\pi, \mu) \right\}, \tag{2.4}$$

where $c(\pi, \mu)$ denotes, with abuse of notation, the information cost induced by the action distribution π.[16] Lemma 2 in the online appendix of Matějka and McKay (2015) implies that the agent's problem has a unique solution in either formulation (up to signal labels). Let $\beta(a_i)$ denote the respective unconditional probability of choosing alternative a_i (calculated using the agent's own prior belief μ):

$$\beta(a) \equiv \sum_{\omega \in \Omega} \mu(\omega) \pi(a|\omega). \tag{2.5}$$

The principal's problem can then be rewritten as choosing $\mu \in \mathcal{M}$ that solves

$$\begin{aligned} &\max_{\mu} \left\{ \sum_{\omega \in \Omega} \mu_p(\omega) \left(\sum_{a \in \mathcal{A}} \pi_\mu(a|\omega) u(a, \omega) \right) \right\} \\ &\text{s.t. } \pi_\mu \text{ solves (2.4) given } \mu. \end{aligned} \tag{2.6}$$

In what follows, we refer to problem (2.6) as the principal's **full problem**. Our main

[16] Cost $c(\pi, \mu)$ is calculated as the cost $c(\phi, \mu)$ of the cheapest strategy (ϕ, σ) that generates π. The choice rule in such a strategy is deterministic, and the signal strategy prescribes at most one signal per action (Matějka and McKay, 2015, Lemma 1). Given this, we have that

$$c(\pi, \mu) = \lambda \left(\sum_{\omega \in \Omega} \mu(\omega) \left(\sum_{a \in \mathcal{A}} \pi(a|\omega) \ln \pi(a|\omega) \right) - \sum_{a \in \mathcal{A}} \beta(a) \ln \beta(a) \right).$$

interest in what follows lies in the properties of the solution μ^* of the full problem and the chosen agent's optimal strategy π_{μ^*}.

We now proceed to analyze the model described above.

2.4 Binary Case

We start by looking at the binary-state, binary-action version of the model, since the results can be articulated more clearly in such a setting. This version differs from the example in Section 2.2 in that the agent has flexibility in his learning technology (as opposed to binary all-or-nothing learning). We show that with the entropy cost function, the principal has to balance off the amount of information acquired against the *nature* of information acquired – since agents with different prior beliefs bias their learning towards different states. This makes the principal favor agents who are somewhat more uncertain than her regarding the state, but who do not necessarily have a uniform prior belief (Proposition 2.2).

To remind, Section 2.2 assumed that the state space is $\Omega = \{l, r\}$, the action set is $\mathcal{A} = \{L, R\}$, and the common utility function net of information costs is such that $u(L|l) = u(R|r) = 1$ and $u(L|r) = u(R|l) = 0$. We proceed by the backward induction, looking at the agent's problem first, and then using the agent's optimal behavior to solve the principal's problem of choosing the best agent.

The agent is allowed to choose any informational structure (Blackwell experiment) he wants, paying the cost which is proportional to the expected reduction of the Shannon entropy of his belief. Using the result presented in Section 2.3.5, the agent's problem can be reformulated as the problem of choosing a stochastic decision rule $\pi : \Omega \to \Delta(\mathcal{A})$, which solves

$$\max_{\pi} \Big\{ \mu\pi(R|r) + (1-\mu)\pi(L|l) - c(\pi, \mu) \Big\}. \tag{2.7}$$

The solution to this problem can be summarized by the two precisions $\{\pi(R|r), \pi(L|l)\}$ or, alternatively, the two unconditional probabilities $\{\beta(R), \beta(L)\}$. Using Theorem 1 in Matějka and McKay (2015), we get that

$$\pi(L|l) = \frac{\beta(L)e^{\frac{1}{\lambda}}}{\beta(L)e^{\frac{1}{\lambda}} + \beta(R)}, \qquad \pi(R|r) = \frac{\beta(R)e^{\frac{1}{\lambda}}}{\beta(L)e^{\frac{1}{\lambda}} + \beta(R)}, \tag{2.8}$$

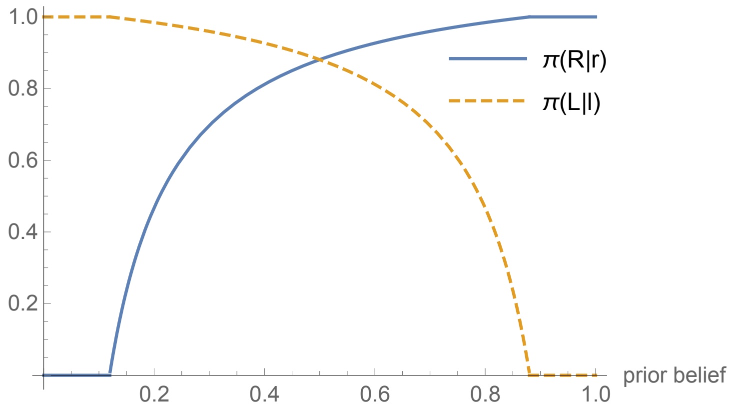

Figure 2.1: Solution of problem (2.7) for different prior beliefs μ.

and their Corollary 2 then adds the conditions

$$\frac{\mu}{\beta(L)+\beta(R)e^{\frac{1}{\lambda}}}+\frac{(1-\mu)e^{\frac{1}{\lambda}}}{\beta(L)e^{\frac{1}{\lambda}}+\beta(R)}=1, \tag{2.9}$$

$$\frac{\mu e^{\frac{1}{\lambda}}}{\beta(L)+\beta(R)e^{\frac{1}{\lambda}}}+\frac{(1-\mu)}{\beta(L)e^{\frac{1}{\lambda}}+\beta(R)}=1. \tag{2.10}$$

Combining (2.8)–(2.10), we get that the solution to problem (2.7) is given by[17]

$$\begin{aligned}\pi(R|r) &= \frac{\left(\mu e^{\frac{1}{\lambda}}-(1-\mu)\right)e^{\frac{1}{\lambda}}}{\left(e^{\frac{2}{\lambda}}-1\right)\mu},\\ \pi(L|l) &= \frac{\left((1-\mu)e^{\frac{1}{\lambda}}-\mu\right)e^{\frac{1}{\lambda}}}{\left(e^{\frac{2}{\lambda}}-1\right)(1-\mu)},\end{aligned} \tag{2.11}$$

cropped to $[0,1]$. Figure 2.1 demonstrates how the agent's action precisions choice depends on his prior belief.

In turn, the principal's problem is the same as in Section 2.2:

$$\begin{gathered}\max_{\mu}\left\{\mu_p\pi_\mu(R|r)+(1-\mu_p)\pi_\mu(L|l)\right\}\\ \text{s.t. } \pi_\mu \text{ solves problem (2.7) given } \mu.\end{gathered} \tag{2.12}$$

It is easy to see by comparing the payoffs in (2.7) and (2.12) that the principal benefits

[17]This solution takes the form of the so-called rational inattention (RI) logit. In comparison to the standard logit behavior, under RI-logit the decision-maker (the agent in our case) has a stronger tendency to select the ex ante optimal alternatives more frequently.

from higher precisions $\pi(R|r)$ and $\pi(L|l)$, the same as the agent. However, the relative weights the principal and the agent assign to these precisions depend on their respective priors μ_p and μ, and hence differ between the two. Hence, in order to understand the trade-offs that the principal faces in hiring agents with different priors, we need to explore how the agent's optimal strategy (2.11) depends on his prior belief μ.

Solving the problem (2.7), the agent faces a trade-off between increasing the precision of his decisions, $\pi(R|r)$ and $\pi(L|l)$, and the cost of information. With the flexible acquisition technology, the agent prefers to learn more about the more probable event. Namely, the higher is the probability that the agent's prior belief assigns to $\omega = r$, the more important is precision $\pi(R|r)$ for his payoff, compared to $\pi(L|l)$. Therefore, two agents with different beliefs would acquire *different* information, leading to different precisions $\pi(R|r)$ and $\pi(L|l)$.[18] At the same time, the closer is the prior belief μ to the extremes ($\mu = 0$ or $\mu = 1$), the more confident is the agent about what the state is, and the less relevant is the precision in the other state for him, leading to such an agent acquiring less information in total.

To summarize, the agent's belief μ affects his optimal decision precisions in two ways: a more uncertain agent acquires more information (and hence makes a better decision on average) than an agent who believes one state is more likely. However, the latter is more concerned with choosing the correct action in the ex ante more likely state, while neglecting the other state.

The principal prefers, ceteris paribus, to hire an agent who acquires more information and hence makes better choices – i.e., a more uncertain agent (μ close to 0.5). However, if she believes that, e.g., state r is ex ante more likely ($\mu_p > 0.5$), then she, for all the same reasons as the agent, cares more about the agent choosing the optimal action in state r than in state l. The latter leads her to prefer an agent who is not completely uncertain ($\mu \neq 0.5$), favoring those who agree with her in terms of which state is more likely ($\mu > 0.5$). Balancing the two issues leads to the principal optimally hiring an agent who has a belief different from hers: $\mu \neq \mu_p$, yet who fundamentally agrees with her on the ex ante optimal action: $\mu \geq 0.5 \iff \mu_p \geq 0.5$.

Figure 2.2 plots the principal's expected utility from hiring an agent as a function of the agent's belief μ when $\mu_p = 0.7$. We can see the principal with a prior belief $\mu_p = 0.7$

[18] This feature of the flexible information acquisition model was analyzed in the application to belief polarization by Nimark and Sundaresan (2019), as well as in the marketing literature (see Jerath and Ren [2021]).

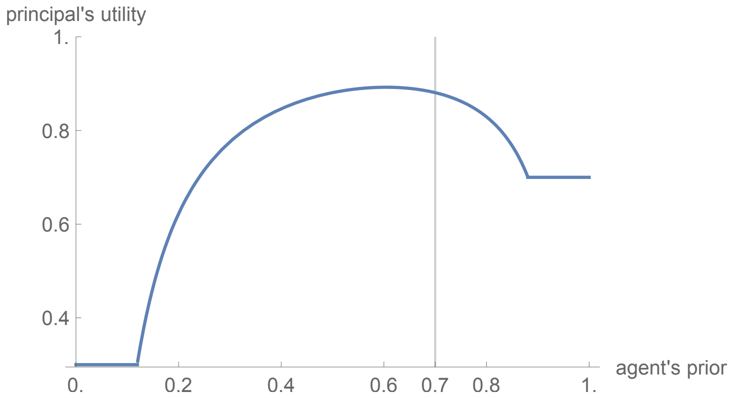

Figure 2.2: Expected utility of principal with prior belief $\mu_p = 0.7$ as a function of the agent's prior belief μ.

would prefer to hire an agent with a prior belief $\mu \approx 0.6$. Note that the graph is flat for very high and very low μ, which corresponds to the agents who do not learn anything, and simply always choose the ex ante optimal action. Further, agents with low $\mu \sim (0.15, 0.2)$ acquire non-trivial information, but hiring them is worse for the principal than taking the ex ante optimal action (equivalent to hiring an agent with $\mu = 1$). In other words, if an agent is too biased, the information he acquires does not benefit the principal due to the bias in the agent's actions introduced by the agent's prior belief being also biased (from the principal's standpoint).

Proposition 2.2 below formalizes this intuition and provides a closed-form solution for the optimal delegation strategy given the principal's prior belief μ_p. Figure 2.3 visualizes the optimal delegation strategy as a function of μ_p.

Proposition 2.2. *If $\mathcal{M} = [0, 1]$, then the principal's optimal delegation strategy is given by*

$$\mu^* = \frac{\sqrt{\mu_p}}{\sqrt{\mu_p} + \sqrt{1 - \mu_p}}. \tag{2.13}$$

Therefore, if $\mu_p \in \left(\frac{1}{2}, 1\right)$, the principal optimally delegates to an agent with belief $\mu_A \in \left(\frac{1}{2}, \mu_p\right)$.

One thing to note about Proposition 2.2 is that the optimal delegation strategy (2.13) does not depend on the agent's information cost factor, λ. While it is immediate that the higher is λ, the less information the agent with any given prior μ collects, Proposition 2.2 serves to show that the trade-off between the quantity of information and the bias in the decisions does not depend on the absolute quantity of information the agent acquires.

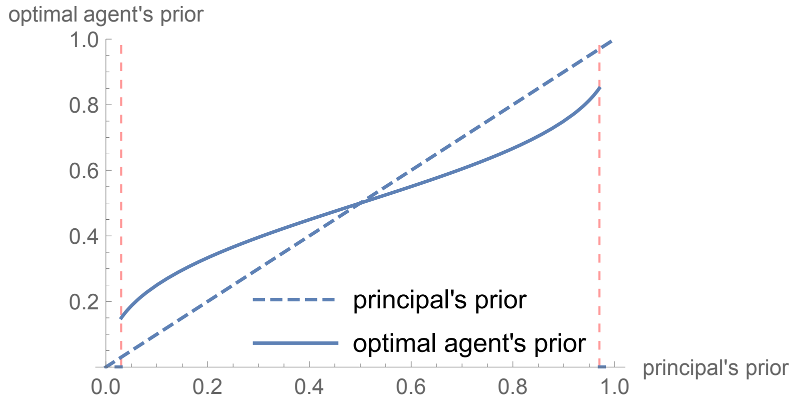

Figure 2.3: The optimal delegation strategy μ^* as a function of the principal's prior belief μ_p.

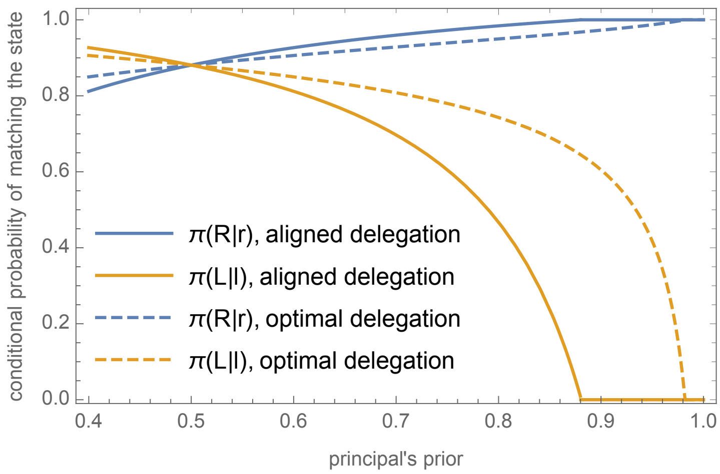

Figure 2.4: Action precisions under optimal delegation and delegating to the aligned agent.

Figure 2.4 demonstrates the difference in the action precisions between delegating to a perfectly aligned agent ($\mu = \mu_p$) and the optimally misaligned agent as given by (2.13). Optimal delegation leads to the agent consuming more information, lowers the probability of correctly matching the ex ante more likely (according to the principal's belief μ_p) state, $\pi(R|r)$, and increases $\pi(L|l)$, thereby bringing the two closer together. Overall, under the optimal delegation, the ex ante less attractive option (as seen by both the principal and the agent) is implemented relatively more frequently as compared to the case of the aligned delegation. The principal's benefit from a higher $\pi(L|l)$ under optimal delegation outweighs her loss from a lower $\pi(R|r)$ than under aligned delegation.

Here, an interesting connection can be made to prospect theory (see Barberis [2013] for a review). In particular, Tversky and Kahneman (1992) suggest that in problems of choice under risk, individual decision-makers tend to succumb to cognitive biases such as overweighing small probabilities and underweighing large probabilities. They propose a probability weighting function that decision-makers unconsciously use, which is reminiscent of our optimal delegation strategy (2.13), with μ_p being the objective probability and μ^* being the decision-maker's perceived probability. Our result can thus be interpreted as one possible evolutionary explanation of the probability weighting functions. Namely, suppose that "Nature" (evolutionary pressure) is the principal and "Human" is the agent. They both have common utility function $u(a, \omega)$ representing the survival probability of the individual/population, but natural selection is indifferent towards the human's cognitive costs $c(\phi, \mu)$ involved in the decision-making process. In this setting, natural selection would lead humans to develop probabilistic misperceptions according to (2.13), since these maximize the survival probability.[19]

In the next section, we generalize the binary model, assuming more available alternatives, while keeping the structure of the payoffs the same.

2.5 General Case

In this section, we extend the analysis to a general problem of finding the best alternative, allowing for $N > 2$ actions and states. We show that the principal's optimal delegation strategy is qualitatively the same as in the binary case, i.e., it is optimal to hire a "more uncertain" agent who investigates more actions in search of the best one than a fully aligned agent. Further, we characterize the whole set of decision rules that can be achieved by selecting the agent's prior belief and show that it coincides with what can be achieved by selecting action-contingent subsidies for the agent.

We are now looking at the model with $\mathcal{A} \equiv \{a_1, ..., a_N\}$ and $\Omega \equiv \{\omega_1, ..., \omega_N\}$ for some N, and the preferences are given by $u(a_i, \omega_i) = 1$ and $u(a_i, \omega_j) = 0$ if $i \neq j$. Without loss of generality, we assume that the principal's belief μ_p is such that $\mu_p(\omega_1) \geq \mu_p(\omega_2) \geq \ldots \geq \mu_p(\omega_N)$ (otherwise relabel states and actions as necessary). As before, the results from Section 2.3.5 apply, meaning that the agent's problem is equivalent to choosing the action distribution $\pi : \Omega \to \Delta(\mathcal{A})$ to maximize (2.4), and the principal selects an agent according to his prior $\mu \in \mathcal{M}$ to maximize (2.6). We do not restrict the choice of agents

[19]Steiner and Stewart (2016) suggest an alternative explanation of probabilistic misperceptions using a similar nature-as-a-principal approach, but a different source of conflict between Nature and Human.

and let $\mathcal{M} = \Delta(\Omega)$ (i.e., for any probability distribution $\mu \in \Delta(\Omega)$, the principal can find and hire an agent with prior belief μ).

2.5.1 Agent's Problem

Proceeding by backward induction, we start by looking at the problem of an agent with some prior belief μ. Invoking Theorem 1 from Matějka and McKay (2015), as we did in the binary case, we obtain that the agent's optimal decision rule satisfies:

$$\pi(a_i|\omega_j) = \frac{\beta(a_i)e^{\frac{u(a_i,\omega_j)}{\lambda}}}{\sum_{k=1}^{N}\beta(a_k)e^{\frac{u(a_k,\omega_j)}{\lambda}}}, \tag{2.14}$$

where $\beta(a_i)$, defined in (2.5), is the unconditional choice probability according to the agent's prior belief μ, and itself depends on $\{\pi(a_i|\omega_j)\}_{j=1}^{N}$. While (2.14) does not provide a closed-form solution for the decision rule $\pi(a_i|\omega_j)$, it implies that the conditional choice probabilities π are unambiguously determined given the unconditional choice probabilities β, and this mapping depends solely on the agent's payoffs and not on his prior belief. In what follows, we use the implication that a collection of the unconditional choice probabilities β pins down the whole decision rule π and use β to summarize the agent's chosen decision rule.

The above is not to say that closed-form expressions cannot be obtained. Caplin et al. (2019) show (see their Theorem 1) that an agent with a prior belief μ optimally chooses a decision rule that generates unconditional choice probabilities

$$\beta(a_i) = \max\left\{0, \frac{1}{\delta}\left(\frac{(K(\beta)+\delta)\mu(\omega_i)}{\sum\limits_{j\in C(\beta)}\mu(\omega_j)} - 1\right)\right\}, \tag{2.15}$$

where $C(\beta) \equiv \{i \in \{1, ..., N\} : \beta(a_i) > 0\}$ denotes the **consideration set**, i.e., the set of actions that are chosen with strictly positive probabilities, and $K(\beta) \equiv |C(\beta)|$ denotes the power (number of actions in) this set.

2.5.2 Principal's Relaxed Problem

As mentioned previously, (2.14) implies that a collection of the unconditional choice probabilities β pins down the whole decision rule π. Let us then consider a **relaxed problem** for the principal, in which instead of choosing the agent's prior μ, she is free to select the unconditional choice probabilities $\beta \in \Delta(\mathcal{A})$ directly:

$$\max_{\beta} \left\{ \sum_{j=1}^{N} \mu_p(\omega_j) \left(\sum_{i=1}^{N} \frac{\beta(a_i) e^{\frac{u(a_i,\omega_j)}{\lambda}}}{\sum_{k=1}^{N} \beta(a_k) e^{\frac{u(a_k,\omega_j)}{\lambda}}} u(a_i,\omega_j) \right) \right\}. \tag{2.16}$$

In the above, we used (2.14) to represent the conditional probabilities $\pi(a_i|\omega_j)$ in (2.6) in terms of the unconditional probabilities $\beta(a_i)$. In Section 2.5.3 we show that the solution to this relaxed problem is implementable in the full problem – i.e., that there exists an agent's belief μ that generates the principal-optimal choice probabilities β.

Note that $\beta(a_i)$ in the above represents the probability with which *an agent* expects to select action a_i. The principal's expected probability of a_i being selected, $\sum_{j=1}^{N} \mu_p(\omega_j)\pi(a_i|\omega_j)$, would generically be different, since her prior belief μ_p is different. Analyzing the principal's problem through the prism of choosing β is the most convenient approach due to the RI-logit structure of the solution to the agent's problem.

Given the state-matching preferences $u(a_j,\omega_j) = 1$, $u(a_i,\omega_j) = 0$ if $i \neq j$, we can simplify (2.16) to

$$\max_{\beta} \left\{ \sum_{j=1}^{N} \mu_p(\omega_j) \frac{\beta(a_j) e^{\frac{1}{\lambda}}}{1 + \delta\beta(a_j)} \right\}, \tag{2.17}$$

where $\delta \equiv e^{\frac{1}{\lambda}} - 1$. We can now state the solution to the principal's problem as follows.

Lemma 2.1. *The solution to the principal's relaxed problem* (2.17) *is given by*

$$\beta^*(a_i) = \max \left\{ 0, \frac{1}{\delta} \left(\frac{(K(\beta^*) + \delta)\sqrt{\mu_p(\omega_i)}}{\sum\limits_{j \in C(\beta^*)} \sqrt{\mu_p(\omega_j)}} - 1 \right) \right\},$$

where $\delta \equiv e^{\frac{1}{\lambda}} - 1$.

We explore the solution and the intuition behind it in terms of the action choice probabilities in Section 2.5.4. Before that, however, we need to ensure that this solution is attainable in the principal's full problem, which is done in the following section.

2.5.3 Principal's Full Problem

The question this section explores is: can the principal generate choice probabilities β^* by appropriately choosing the agent's prior belief μ? In the binary case, the answer was trivially "yes": due to continuity of the agent's strategy, by varying the agent's belief $\mu(r)$

between 0 and 1, the principal could induce any unconditional probability $\beta(R)$. In the multidimensional case, this is not immediate. However, the following theorem shows that the result still holds with N actions and states under state-matching preferences.

Theorem 2.1. *In the principal's full problem* (2.6)*, any vector* $\beta \in \Delta(\mathcal{A})$ *of unconditional choice probabilities is* ***implementable****: there exists a prior belief* $\mu \in \Delta(\Omega)$ *such that* $\beta(a_i) = \sum_{j=1}^{N} \mu(\omega_j)\pi_\mu^*(a_i|\omega_j)$*, where* π_μ^* *solves the agent's problem* (2.4) *given* μ*.*

The theorem states that if $\mathcal{M} = \Delta(\Omega)$, then the principal can generate any vector of unconditional action probabilities. Note that this does not imply that she is able to select any decision rule $\pi(a_i|\omega_j)$ – if this were the case, under the state-matching preferences she would simply choose to have $\pi(a_i|\omega_i) = 1$ for all i. However, Theorem 2.1 does imply that the choice probabilities described in Lemma 2.1 – those that solve the principal's relaxed problem, – are implementable and thus also solve her full problem.

The result does, however, rely on the state-matching preferences: we show in Section 2.6.1 that it does not hold for arbitrary payoffs.

2.5.4 Properties of the Optimal Delegation Strategy

While Lemma 2.1 presents the solution of the principal's problem in terms of the unconditional choice probabilities, this representation is not the most visual. We now demonstrate some implications of this solution in terms of other variables. Namely, Proposition 2.3 extends Proposition 2.2 and shows how the chosen agent's prior belief relates to that of the principal. Proposition 2.4 then compares actions taken under optimal delegation vs aligned delegation.

We begin by looking at the optimal agent choice in terms of the agent's belief μ^*.

Proposition 2.3. *The principal's equilibrium delegation strategy* μ^* *is such that for all* $i, j \in \{1, ..., N\}$*:*

$$\frac{\mu^*(\omega_i)}{\mu^*(\omega_j)} = \frac{\sqrt{\mu_p(\omega_i)}}{\sqrt{\mu_p(\omega_j)}}.$$

In particular, $\mu^*(\omega_1) \geq ... \geq \mu^*(\omega_N)$*. Further,* $\mu^*(\omega_1) \leq \mu_p(\omega_1)$ *and* $\mu^*(\omega_N) \geq \mu_p(\omega_N)$*, with equalities if and only if* $\mu_p(\omega_1) = ... = \mu_p(\omega_j)$*.*

The intuition behind the proposition above is the same as that behind Proposition 2.2: the optimally chosen agent is more uncertain than the principal between any given pair of states. To see this, note that if $\mu_p(\omega_i) > \mu_p(\omega_j)$ then $1 < \frac{\mu^*(\omega_i)}{\mu^*(\omega_j)} < \frac{\mu_p(\omega_i)}{\mu_p(\omega_j)}$ – i.e., the agent believes state ω_i is ex ante more likely than ω_j, as the principal does, but he assigns relatively less weight to ω_i. This applies to any pair of states. Thus, the implication is that the optimal agent must assign a lower ex ante probability to ω_1, the most likely state according to the principal, than she does, and vice versa for ω_N. Note further that the result in Proposition 2.3 is again independent of λ, implying that the optimal delegation strategy is determined by the relative trade-off between the quantity of information acquired and the bias introduced in actions by the misalignment in beliefs, but the absolute quantity of information acquired is irrelevant. In particular, hiring an agent with μ^* is optimal even when he acquires no information, and another agent μ is available, who would be willing to invest effort in learning ω (since such a μ-agent would be too biased relative to the principal).

We now switch to comparing the choices made under optimal delegation to those that would arise under **aligned delegation** – i.e., if the principal selected an agent with $\mu = \mu_p$. Let $\bar{\beta}$ denote the choice probabilities that would be generated under aligned delegation. Caplin et al. (2019) show that these probabilities $\bar{\beta}$, as a function of the agent's prior μ, are given by (see their Theorem 1)

$$\bar{\beta}(a_i) = \max\left\{0, \frac{1}{\delta}\left(\frac{(K(\bar{\beta})+\delta)\mu(\omega_i)}{\sum\limits_{j\in C(\bar{\beta})}\mu(\omega_j)} - 1\right)\right\}. \tag{2.18}$$

Since $\mu_p(\omega_1) > ... > \mu_p(\omega_N)$, the consideration set in the aligned problem is then simply $C(\bar{\beta}) = \{1, ..., \bar{K}\}$, and its size $\bar{K} \equiv K(\bar{\beta})$ is the unique solution of

$$\mu_p(\omega_{\bar{K}}) > \frac{1}{\bar{K}+\delta}\sum_{j=1}^{\bar{K}}\mu_p(\omega_j) \geq \mu_p(\omega_{\bar{K}+1}). \tag{2.19}$$

In turn, we can see from Lemma 2.1 that under optimal delegation, size $K^* = K(\beta^*)$ of the consideration set under optimal choice is

$$\sqrt{\mu_p(\omega_{K^*})} > \frac{1}{K^*+\delta}\sum_{j=1}^{K^*}\sqrt{\mu_p(\omega_j)} \geq \sqrt{\mu_p(\omega_{K^*+1})}. \tag{2.20}$$

These two conditions allow us to compare K^* and $\bar{K}$ directly, which is done by the following proposition.

Proposition 2.4. *Optimal delegation weakly expands the consideration set relative to aligned delegation:*

$$K(\beta^*) \geq K(\bar{\beta}).$$

In other words, delegating to an optimally misaligned agent leads to a wider variety of actions played in equilibrium. This is a direct consequence of delegation to a more uncertain agent – since he is less sure than the principal of what the optimal action is ex ante, he considers more actions worth investigating. Every action has some positive probability of actually being optimal, and thus a more uncertain agent plays a wider range of different actions ex post. We could already see this effect at play in the binary case, where if μ_p is extreme, then an aligned agent takes the ex ante optimal action without acquiring any additional information, whereas the optimally chosen agent could investigate both actions.

2.6 Misaligned Beliefs Versus Other Instruments

The preceding analysis above explored the problem of selecting an agent according to their prior belief. It provided grounds for using misaligned beliefs as an instrument in delegation, but it is still worth studying how this instrument compares to the other instruments, such as contracting or restricting the delegation set. In this section, we draw this comparison. We keep the overall structure of the problem the same as in Section 2.5, but now we consider three separate versions in which the principal has different tools at her disposal, and we compare outcomes to those in the baseline problem of choosing the agent's beliefs.

2.6.1 Contracting on Actions/Misaligned Preferences

The most basic delegation tool is contracting: if the principal could offer the agent a contract that specifies contingent payments, this would be the most direct way to provide incentives.[20] We can think of two main options here: contracting on outcomes (where "outcome" is understood in the sense of "was the agent's action correct?") and contracting on actions. The former requires that both outcomes are contractible (i.e., observable and

[20] See Laffont and Martimort (2009) for many examples.

verifiable), the latter imposes such requirement on actions. Both options require that the principal has the freedom to design the payments, which is a strong assumption in itself.

We begin by exploring contracting on actions in our framework, in which the principal must design a payment schedule $\tau : \mathcal{A} \to \mathbb{R}$ to be paid to the agent. We assume that all agents and the principal have a common prior belief μ_p, all players' preferences are quasilinear in payments, and the principal's marginal utility of money is ρ, and the agent's marginal utility of money is 1. For some results, we additionally impose the limited liability assumption ($\tau(a_i) \geq 0$ for all i).[21]

Note that instead of contracting, we can interpret this setup as a problem of selecting an agent with misaligned preferences by setting $\rho = 0$. Schedule τ then represents an agent's "biases", i.e., inherent preferences towards certain actions on top of the "unbiased" utility $u(a, \omega)$. Such a problem of selecting an agent with optimally misaligned preferences is a natural counterpart to our baseline problem of selecting an agent with optimally misaligned beliefs.

The agent's problem (again using the equivalence presented in 2.3.5) is then given by

$$\max_{\pi} \left\{ \sum_{j=1}^{N} \mu_p(\omega_j) \sum_{i=1}^{N} \pi(a_i|\omega_j)\big(u(a_i, \omega_j) + \tau(a_i)\big) - c(\phi, \mu_p) \right\}, \tag{2.21}$$

given τ, and the principal's **contracting problem** is

$$\max_{\tau} \left\{ \sum_{j=1}^{N} \mu_p(\omega_j) \sum_{i=1}^{N} \pi(a_i|\omega_j)\big(u(a_i, \omega_j) - \rho\tau(a_i)\big) \right\}, \tag{2.22}$$

subject to π solving (2.21) given τ.

Instead of providing a closed-form solution to this problem, we appeal to Theorem 2.1 to argue that regardless of ρ, the principal cannot obtain higher expected utility than in the baseline problem of choosing an agent with a misaligned belief μ. In particular, Theorem 2.1 implies that any unconditional choice probabilities $\beta \in \Delta(\mathcal{A})$ generated by an agent, who is incentivized by payments or misaligned preferences, can also be obtained by selecting an agent with appropriately misaligned beliefs. Moreover, if we relax limited liability and allow negative payments, then by using the results of

[21] In line with the baseline problem, we do not impose any participation constraints on the agent. The implicit assumption here is that the agent is being paid some non-negotiable unconditional salary if he is hired, which is sufficient to ensure participation. Payments $\{\tau(a_i)\}$ should then be treated as premia, with the limited liability assumption implying they must be non-negative.

Matveenko and Mikhalishchev (2021) we can also show the converse – that any decision rule achievable with misaligned beliefs can be replicated with payments τ (or by setting the quotas – i.e., imposing specific unconditional choice probabilities for a different action). These results are formalized in the following proposition.[22]

Proposition 2.5. *The principal's problem of contracting on actions* (2.22) *is equivalent to her full (delegation) problem* (2.6)*:*

1. *For any vector $\tau : \mathcal{A} \to \mathbb{R}$ of payments/biases and a corresponding $\beta : \Omega \to \Delta(\mathcal{A})$ that solves* (2.21) *given τ, there exists a prior belief $\mu \in \Delta(\Omega)$ such that β also solves* (2.4) *given μ.*

2. *For any $\mu \in \Delta(\Omega)$ and the corresponding $\beta : \Omega \to \Delta(\mathcal{A})$ that solves* (2.4) *given μ, there exist payments $\tau : \mathcal{A} \to \mathbb{R}$ such that β also solves* (2.21) *given τ.*

The proposition above directly implies that neither of the two instruments (contracting on actions or searching for an agent with stronger/weaker preferences for specific actions) can yield better results than hiring an agent with an optimally misaligned belief. If limited liability is in place ($\tau(a_i) \geq 0$ for all i), then contracting on actions is strictly worse, since it cannot yield a better decision rule, but requires payments from the principal – payments which are avoidable if she instead hires an agent who is intrinsically motivated by his beliefs over states or preferences towards specific actions.

Further, Theorem 2.1 and the results of Matveenko and Mikhalishchev (2021) also imply that if we fix $u(a, \omega_i) - u(a, \omega_j)$ for all a, ω_i, ω_j and vary the rest of the agent's preferences and beliefs arbitrarily, the resulting set of distributions $\beta \in \Delta(\mathcal{A})$ can be fully covered by only varying one of the two. This implies that no combination of misaligned beliefs, misaligned preferences, and payments for actions can perform better than any individual instrument. Moreover, it also implies that suboptimal misalignment along any dimension can be amended using other instruments. That is, if a given agent holds a non-optimal prior belief (that does not coincide with the principal's either), the optimal behavior might be induced via action-contingent transfers. Conversely, if all agents are biased towards certain actions, this misalignment can be compensated for by selecting an agent with the right prior belief. The following proposition presents one example of such equivalence, in the context of a model with $N = 2$.

[22] The result regarding quotas is not included in the proposition, yet it follows immediately from Lemma 1 of Matveenko and Mikhalishchev (2021).

Proposition 2.6. *Consider the binary setting of Section 2.4. Consider the principal's problem of contracting on actions* (2.22)*, where* $\rho = 0$ *and the agent holds prior belief* $\mu \neq \mu_p$*. Then:*

1. *for any* μ*, there exist payments/biases* $\{\tau^*(L), \tau^*(R)\}$ *that implement the optimal conditional choice probabilities from Section 2.4;*

2. *these payments/biases are such that*[23]

$$\tau^*(R) \geq \tau^*(L) \iff \mu \leq \mu^* = \frac{\sqrt{\mu_p}}{\sqrt{\mu_p} + \sqrt{1-\mu_p}}.$$

It is easy to see the intuition behind the proposition: if the agent's prior belief μ assigns lower probability to state $\omega = r$ compared to the principal-optimal prior μ^* given in Proposition 2.2, such an agent is ex ante too biased towards action L for the principal's taste, even though he potentially acquires more information than an agent with belief μ^*. Therefore, the principal can nudge the agent towards action R by offering higher payment if he selects R (or find an agent whose preference bias towards R offsets his belief bias towards state l).

2.6.2 Contracting on Outcomes

We now turn to contracting on outcomes. The outcome in our model is effectively binary: whether a correct action was chosen ($a = a_j$ when $\omega = \omega_j$) or not. We thus let the principal select payments $\bar{\tau}, \underline{\tau}$ that the agent receives, so that $\tau(a_i, \omega_i) = \bar{\tau}$ and $\tau(a_i, \omega_j) = \underline{\tau}$ if $i \neq j$.[24] We again assume limited liability ($\bar{\tau}, \underline{\tau} \geq 0$), preferences that are quasilinear in payments for all agents, and let the agent's marginal utility of money to be 1, and the principal's marginal utility of money to be ρ.

The agent's problem is then choosing $\pi : \Omega \to \Delta(\mathcal{A})$ that solves[25]

$$\max_{\pi} \left\{ \sum_{j=1}^{N} \mu(\omega_j) \sum_{i=1}^{N} \pi(a_i|\omega_j)\big(u(a_i,\omega_j) + \tau(a_i,\omega_j)\big) - c(\phi,\mu) \right\}, \tag{2.23}$$

[23] *The closed-form expressions are available in the proof in the Appendix.*

[24] If the principal could contract on both actions and outcomes, she would have the freedom to select any payment schedule $\{\tau(a_i, \omega_j)\}$. Lindbeck and Weibull (2020) study such a problem with N states and two actions.

[25] While it is more common in the literature to consider an agent who yields no intrinsic utility from actions and is motivated exclusively via payments, for sake of consistency, we maintain the assumption that the agent enjoys the same intrinsic utility $u(a,\omega)$ as the principal, albeit possibly to a different magnitude.

and the principal's **contracting problem** is

$$\max_{\bar{\tau},\underline{\tau}} \left\{ \sum_{j=1}^{N} \mu_p(\omega_j) \sum_{i=1}^{N} \pi(a_i|\omega_j)\big(u(a_i,\omega_j) - \rho\tau(a_i,\omega_j)\big) \right\}, \tag{2.24}$$
$$\text{s.t. } \tau(a_i,\omega_i) = \bar{\tau} \text{ for all } i,$$
$$\tau(a_i,\omega_j) = \underline{\tau} \text{ for all } i, j \neq i,$$

and subject to β corresponding to a solution of (2.23) given $\bar{\tau}, \underline{\tau}$.

It is trivially optimal for the principal to set $\underline{\tau} = 0$, since her objective is to provide incentives for the agent to match the state. Then, however, the agent's (ex post) payoff net of information cost becomes $u(a_i,\omega_j)+\tau(a_i,\omega_j) = (1+\bar{\tau})u(a_i,\omega_j)$, and the principal's payoff is $u(a_i,\omega_j)-\tau(a_i,\omega_j) = (1-\rho\bar{\tau})u(a_i,\omega_j)$. In other words, by increasing the incentive payment $\bar{\tau}$, the principal effectively lowers the relative cost of information for the agent, at the cost of decreasing her own payoff. It then appears like an instrument that could be universally useful for the principal – even when she chooses an agent with the optimal prior belief, she could still benefit from reducing the agent's information cost, which would result in him acquiring more information. The following proposition shows, however, that this is not the case: while contracting on outcomes may be a useful instrument, it cannot improve on delegating to the optimally biased agent when payments are costly to the principal.

Proposition 2.7. *Consider the principal's contracting problem* (2.24) *in the binary setting of Section 2.4 and suppose $\rho = 1$ and $\mu_p > 1/2$. Then there exist $\bar{\mu}_1, \bar{\mu}_2$: $\bar{\mu}_1 < \mu^* < \mu_p < \bar{\mu}_2$ (where μ^* is as in Proposition 2.2), such that if either $\frac{1}{1+e^{\frac{1}{\lambda}}} \leq \mu \leq \bar{\mu}_1$, or $\bar{\mu}_2 \leq \mu \leq \frac{e^{\frac{1}{\lambda}}}{1+e^{\frac{1}{\lambda}}}$, then $\bar{\tau} > 0$ solves* (2.24).

The proposition states that the principal uses the incentive payments, $\bar{\tau} > 0$, when there is an intermediate degree of misalignment in opinions with the agent. If the agent has a very extreme prior belief and acquires no information on his own, it may be too costly for the principal to incentivize such an agent to acquire any information. The requirements that $\frac{1}{1+e^{\frac{1}{\lambda}}} \leq \mu \leq \frac{e^{\frac{1}{\lambda}}}{1+e^{\frac{1}{\lambda}}}$ ensure that the agent acquires information voluntarily so that this problem does not arise.[26] On the other hand, if there is too little

[26]Note that there may still be μ outside of this interval for which offering positive incentive payments is optimal for the principal. In that sense, Proposition 2.7 only provides a condition on μ that is sufficient for $\bar{\tau} > 0$. A necessary and sufficient condition would look similarly, but feature different outer boundaries for μ.

misalignment ($\bar{\mu}_1 < \mu < \bar{\mu}_2$), then the agent's information acquisition choice is already in line with the principal's wishes, and any further payments would not have a significant enough effect on the agent's incentives to justify the cost they incur on the principal.

2.6.3 Restricting the Delegation Set

Another instrument commonly explored in the delegation literature is restricting the delegation set – i.e., the set of actions that the agent may take (see, e.g., Holmström [1980]). In particular, in the context of "delegated expertise" problems, Szalay (2005) and Ball and Gao (2021) show that it may be optimal to rule out an ex ante optimal action in order to force the agent to exert effort and learn which of the ex post optimal (but ex ante risky) actions is best. Lipnowski et al. (2020) show a similar result in a Bayesian Persuasion setting in which the receiver is rationally inattentive to the sender's message.

In our setting, however, there are no "safe" actions that the principal could rule out, as Propostion 2.4 suggests. Assuming that the principal and the agent hold the same prior belief μ_p, and $\mu_p(\omega_1) > ... > \mu_p(\omega_N)$, action a_1 is the "safest" in the sense of being the most likely ex ante to be optimal. However, it would be trivially suboptimal for the principal to ban a_1 – since, indeed, this is the action that is ex ante most likely to be ex post optimal! In other words, while excluding a_1 from the delegation set would lead the agent to acquire more information, it would also lead to larger ex post losses due to the agent being unable to select action a_1 in cases in which it is optimal to do so. Thus while the general idea of the principal being willing to nudge the agent to acquire more information/information about ex ante suboptimal actions holds true in our setting, restricting the delegation set is not an instrument that lends any value to the principal.

Proposition 2.8 below summarizes this logic. Consider the agent's problem as given by

$$\max_{\pi} \left\{ \sum_{j=1}^{N} \mu_p(\omega_j) \sum_{i=1}^{N} \pi(a_i|\omega_j) u(a_i, \omega_j) - c(\phi, \mu_p) \right\}, \tag{2.25}$$

given $A^* \subseteq \mathcal{A}$ (and the maximization is w.r.t. a mapping $\pi : \Omega \to \Delta(A^*)$), and the principal's **restriction problem**

$$\max_{A^*} \left\{ \sum_{j=1}^{N} \mu_p(\omega_j) \sum_{i=1}^{N} \pi(a_i|\omega_j) u(a_i, \omega_j) \right\}, \tag{2.26}$$

subject to $\pi : \Omega \to \Delta(A^*)$ solving (2.25) given A^*. Then we can state the result as follows.

Proposition 2.8. *The unrestricted delegation set $A^* = \mathcal{A}$ is always a solution to the principal's restriction problem* (2.26)*.*

2.7 Extensions

2.7.1 Alternative Preference Specifications

The analysis in Sections 2.5 and 2.6.1 is heavily reliant on state-matching preferences that we assume are shared by both the principal and the agent(s). It is reasonable to ask whether our conclusions hold under other preference specifications. Since the utility function $u(a, \omega)$ is shared by both the principal and the agent, it is reasonable to generalize one at a time.

We begin by generalizing the principal's utility function $u_p(a, \omega)$ while maintaining the agent's intrinsic preference for matching the state: $u_A(a_i, \omega_i) = 1$, $u_A(a_i, \omega_j) = 0$ if $i \neq j$. Naturally, the specific functional forms of the optimal delegation strategies (such as those presented in Propositions 2.2 and 2.3, and Lemma 2.1) depend on the specific form of the principal's utility function. However, Theorem 2.1 only depends on the agent's utility function, meaning that Proposition 2.5 still holds: any outcome that can be achieved by contracting on actions or hiring an agent with the misaligned intrinsic preferences, can also be achieved by hiring an agent with misaligned beliefs (and vice versa). Meaning that regardless of the principal's objective function, hiring an agent with state-matching preferences and a suitable belief is as good as hiring an agent with aligned prior belief, state-matching preferences, and either some additional preference over actions, or action-contingent payments on top of that.

The above does, however, hinge on the agent having state-matching preferences as a baseline. Once we allow arbitrary preferences for the agent – even if they align with the principal's preferences net of the information cost – the equivalence stated in Proposition 2.5 breaks down. In such a general case, finding an agent with optimally misaligned preferences may yield strictly better results for the principal than hiring an agent with an optimally misaligned belief, and hence contracting on actions may, in principle, yield better results too. An example (for misaligned preferences) is presented in the proof of the following proposition.

Proposition 2.9. *There exists a utility function* $u(a_i, \omega_j)$ *with* $u(a_i, \omega_i) = \max_k \{u(a_k, \omega_i)\}$ *such that the solution to the principal's relaxed problem* (2.17) *cannot be attained as a solution to the full problem* (2.6)*.*

The optimal decision rule is never in the relative interior, since mixing between two options dominates using mixing from three. However, according to Proposition 3 and Lemma 1 from Matveenko and Mikhalishchev (2021), the principal may implement the optimal solution via action-contingent payments/biased preferences.

2.7.2 Communication

In this section, we consider the importance of decision rights in our model with misaligned beliefs. In particular, we juxtapose the *delegation* scheme explored so far, under which the agent has the power to make the final decision, to *communication*, where an agent must instead communicate his findings to the principal, who then chooses the action. A large literature in organizational economics is devoted to comparing delegation and communication in various settings (see Dessein [2002]; Alonso et al. [2008], and Rantakari (2008) for some examples). We show that in our setting, communication performs as well as delegation – i.e., the principal will always find it optimal to follow the agent's recommendation. This is perhaps unsuprising since Holmström (1980) showed that communication is equivalent to restricting the agent's action set, and this latter problem was shown in Section 2.6.3 to be an irrelevant instrument in our setting, as long as the principal can select an agent with the prior belief she prefers.

Although the principal and an agent have the same preferences, it is generally unclear whether it is optimal for the principal to follow a recommendation of an agent due to the misalignment in their beliefs. Namely, since the principal and the agent start from different prior beliefs, the same is true for posteriors: if the principal could observe the information that the agent obtained, her posterior belief would be different from that of the agent. This implies that ex post, the principal could prefer an action different from the agent's choice, and could benefit ex post from overruling the agent's decision if she had the power to do so. However, this would mean that the agent's incentives to acquire information are different from the baseline model, and could lead to the agent acquiring either more or less information than in the baseline, with the principal having some influence over the agent's learning strategy via her final choice rule.[27] We show below that, in the end, none of these effects comes into play, and there exists a communication equilibrium that replicates the delegation equilibrium.

The setup follows the baseline model from Section 2.3, with the exception that the final stage ("agent selecting action $a \in \mathcal{A}$") is replaced by two. First, after observing signal $s \in \mathcal{S}$ generated by his signal structure ϕ, the agent selects a recommendation (message) $\tilde{a} \in \mathcal{A}$ to the principal. After that, the principal observes the recommendation $\tilde{a}$, uses it to update her belief $\mu_p(\omega|\tilde{a})$ about the state of the world, and then selects an action $a \in \mathcal{A}$ that determines both parties' payoffs.[28] The equilibrium of the communication game is then defined as follows.

Definition 2.2 (Communication Equilibrium)**.** An equilibrium of the cheap talk game is characterized by $(\mu^*, \{\phi^*_\mu, \tilde{\sigma}^*_\mu\}_{\mu \in \mathcal{M}}, \sigma^*, \mu_p)$, which consists of the following:

1. the principal's posterior beliefs $\mu_p : \mathcal{A} \to \Delta(\Omega)$ that are consistent with $(\phi^*_\mu, \sigma^*_\mu)$ (i.e., satisfy Bayes' rule on the equilibrium path);

2. the principal's choice rule $\sigma^* : \mathcal{A} \to \mathcal{A}$, which solves the following for every $\tilde{a} \in \mathcal{A}$, given the posterior μ_p:

$$\max_{\sigma(\tilde{a})} \left\{ \sum_{\omega \in \Omega} \mu_p(\omega|\tilde{a}) u(\sigma(\tilde{a})|\omega) \right\};$$

3. a collection of the agents' information acquisition strategies $\phi^*_\mu : \Omega \to \Delta(\mathcal{S})$ and communication strategies $\tilde{\sigma}^*_\mu : \mathcal{S} \to \mathcal{A}$ that solve the following given σ for every $\mu \in \mathcal{M}$:

$$\max_{\phi, \tilde{\sigma}} \left\{ \sum_{\omega \in \Omega} \mu(\omega) \sum_{s \in \mathcal{S}} \phi(s|\omega) u(\sigma(\tilde{\sigma}(s)), \omega) - c(\phi, \mu) \right\};$$

[27] Argenziano et al. (2016) provide one example of how the principal can manipulate the agent's information acquisition incentives under cheap talk communication.

[28] For simplicity, we assume that the principal only observes the recommendation made by the agent, and not the signal he received or the signal structure he requested. These appear to be reasonable, yet substantial assumptions, and the results would be different if we assumed the principal could observe either the learning strategy, or the realized signal.

4. the principal's choice $\mu^* \in \mathcal{M}$ of the agent to whom the task is delegated, which solves the following given $(\phi_\mu^*, \sigma_\mu^*)$, σ^*, and μ_p:

$$\max_{\mu} \left\{ \sum_{\omega\in\Omega} \mu_p(\omega) \sum_{s\in\mathcal{S}} \phi(s|\omega) u(\sigma(\tilde{\sigma}(s)), \omega) \right\}.$$

We can then state the result as follows.

Proposition 2.10. *There exists a communication equilibrium that is outcome-equivalent to the equilibrium of the original game, in the sense that $(\mu^*, \phi_{\mu^*}^*, \tilde{\sigma}_{\mu^*}^*)$ are the same in both equilibria, and σ^* is an identity mapping.*

This result, however, is subject to a few caveats. First, cheap talk models are plagued by equilibrium multiplicity: for any informative equilibrium, there exist equilibria with less informative communication, up to completely uninformative (babbling) equilibria. In our setting, this means that, in addition to the equilibrium outlined in Proposition 2.10 above, there also exists a babbling equilibrium in which the agent acquires no information and makes a random recommendation, and the principal always ignores it and selects the ex ante optimal action.[29] There would also likely exist multiple equilibria of intermediate informativeness – e.g., equilibria with a limited vocabulary, where only some actions $\tilde{\mathcal{A}} \subset \mathcal{A}$ are recommended on the equilibrium path. In practice, this means that, under communication, there is a risk of miscoordination on uninformative equilibria, whereas under delegation the equilibrium is unique. The same force may also work the other way, and there may be equilibria that are preferred by the principal to the delegation equilibrium, that can only be sustained under cheap talk (see Argenziano et al. [2016] for an example of how such equilibria may arise). However, the question of whether such equilibria exist is beyond the scope of this paper.

The second caveat lies in the fact that Proposition 2.10 relies on the state-matching preferences. In our setting (with the exception of Section 2.7.1), any action is either "right" or "wrong", without any degrees of correctness. The misalignment of beliefs across the principal and the agent is thus small enough to not warrant the principal overriding the agent's suggested action. In contrast, in a uniform-quadratic framework, both states and actions lie in a continuum, and the principal's loss is proportional to the distance

[29] If an agent makes uninformed recommendations, it is optimal for the principal to ignore it. If the principal ignores the recommendation, it is optimal for the agent to not acquire any information. Neither agent in this situation can unilaterally deviate to informative communication.

between the realized state and the chosen action. In such a setting, any misalignment (be it in preferences or beliefs) between the principal and the agent would lead to the principal being willing to override the agent's recommendation, leading to the delegation equilibrium being no longer directly sustainable under communication.

2.8 Conclusion

We show that hiring an agent with beliefs that are misaligned with those of the principal can be beneficial for the principal, contrary to popular belief. In particular, if the agent needs to acquire information to make a decision, delegation to an agent who is ex ante more uncertain about what the best action is but shares some of the principal's predispositions is optimal for her. This is mainly due to a more uncertain agent being willing to acquire more information about the state, which enables more efficient actions to be taken. We show that exploiting belief misalignment can be a valid instrument that the principal can use in delegation, which in some settings is on par with or better than state-contingent transfers or restriction of the action set from which the agent can choose.

In the analysis, we use the workhorse rational inattention model for discrete choice, the Shannon model. It allows us to provide a richer demonstration of the consequences of delegation to a biased agent by allowing the agent to acquire information flexibly, which introduces additional bias into the decisions of an agent with misaligned beliefs. We show that misaligned delegation is optimal *despite* the bias introduced by this flexibility. Hence, while the exact trade-offs do, obviously, depend on the particular cost function specification, the main takeaways would persist regardless of the form of the costs of information.

Our analysis focuses on misalignment and away from contracting. A potentially fruitful avenue for further research would be to consider more carefully the contracting problem in a setting in which the principal and the agent have misaligned beliefs and/or preferences, and to investigate how contracting can build upon the inherent benefits of misalignment.

Further, due to the added complexity of rational inattention models, we confine our exploration to a discrete state-matching model, which strays away from the continuous models more commonly used in delegation problems. In a model with continuous actions, the scope for an agent to manifest his bias is much larger, and hence the trade-off between the agent's information acquisition and biased decision-making would again be different.

Exploration of the effects of misalignment in a continuous model of delegated expertise could be an interesting direction for further research as well.

Yet another assumption that may feel excessively strong in our analysis is the common knowledge of all agents' and the principal's prior beliefs. It may be more reasonable to assume that agents are strategic in presenting their viewpoints to the principal, as well as making inferences from the fact that they were chosen for the job. Such signaling concerns could yield an economically meaningful effect, yet we abstract from them completely in our paper. A more careful investigation is in order.

2.A Main Proofs

2.A.1 Proof of Proposition 2.1

Proved in the text preceding the proposition.

2.A.2 Proof of Proposition 2.2

Throughout this proof, we will refer to the delegation rule under consideration,

$$\mu^* = \frac{\sqrt{\mu_p}}{\sqrt{\mu_p} + \sqrt{1-\mu_p}},$$

as the **candidate rule**. It is straightforward that under the candidate rule, if $\mu_p > \frac{1}{2}$ then $\mu^* \in \left(\frac{1}{2}, \mu_p\right)$, since

$$\frac{\mu^*}{1-\mu^*} = \frac{\sqrt{\mu_p}}{\sqrt{1-\mu_p}} < \frac{\mu_p}{1-\mu_p}$$

when $\mu_p > \frac{1}{2}$, so $\mu^* < \mu_p$, and also $\sqrt{\mu_p} > \sqrt{1-\mu_p}$ in that case, so $\mu^* > \frac{1}{2}$. It thus remains to show that the candidate rule is indeed optimal for the principal. While a shorter proof exists that invokes Lemma 2.1 that derives an optimal strategy for the case of N states and actions, we choose to present a more direct, albeit a somewhat longer, proof.

Plugging the solution to the agent's problem (2.11) (assuming this solution is interior for now) into the principal's problem (2.12), we get that the principal's payoff looks as follows:

$$\begin{aligned}
\mu_p \pi(R|r) + (1-\mu_p)\pi(L|l) &= \mu_p \frac{\left(\mu e^{\frac{1}{\lambda}} - (1-\mu)\right) e^{\frac{1}{\lambda}}}{\left(e^{\frac{2}{\lambda}} - 1\right)\mu} + (1-\mu_p)\frac{\left((1-\mu)e^{\frac{1}{\lambda}} - \mu\right) e^{\frac{1}{\lambda}}}{\left(e^{\frac{2}{\lambda}} - 1\right)(1-\mu)} \\
&= \frac{e^{\frac{1}{\lambda}}}{e^{\frac{2}{\lambda}} - 1}\left[\mu_p\left(e^{\frac{1}{\lambda}} - \frac{1-\mu}{\mu}\right) + (1-\mu_p)\left(e^{\frac{1}{\lambda}} - \frac{\mu}{1-\mu}\right)\right] \\
&\propto e^{\frac{1}{\lambda}} - \mu_p\frac{1-\mu}{\mu} - (1-\mu_p)\frac{\mu}{1-\mu}.
\end{aligned}$$

The FOC for the principal's maximization problem above w.r.t. μ is

$$\frac{\mu_p}{\mu^2} - \frac{1-\mu_p}{(1-\mu_p)^2} = 0$$

$$\iff \frac{\mu}{1-\mu} = \frac{\sqrt{\mu_p}}{\sqrt{1-\mu_p}}. \tag{2.27}$$

It is trivial to verify that the second-order condition holds as well, hence as long as (2.27) yields an interior solution (i.e., the probabilities in (2.11) are in $[0,1]$), the candidate solution is indeed optimal among all such interior solutions.

We now check for which μ the solution (2.11) is interior. Using the expressions (2.11), one can easily verify that $\pi(R|r) \geq 0 \iff \frac{\mu}{1-\mu} \geq e^{-\frac{1}{\lambda}}$ and $\pi(R|r) \leq 1 \iff \frac{\mu}{1-\mu} \leq e^{\frac{1}{\lambda}}$, and the conditions $\pi(L|l) \in [0,1]$ yield the same two interiority conditions. This implies that if $\frac{\mu}{1-\mu} \in \left[e^{-\frac{1}{\lambda}}, e^{\frac{1}{\lambda}}\right]$, then the agent acquires some information and selects both actions with positive probabilities, and otherwise $(\pi(R|r), \pi(L|l)) \in \{(1,0),(0,1)\}$, meaning that the agent simply chooses the ex ante optimal action for sure without acquiring any information about the state.

The candidate rule then suggests that the principal delegates to a learning agent iff $\frac{\mu_p}{1-\mu_p} \in \left[e^{-\frac{2}{\lambda}}, e^{\frac{2}{\lambda}}\right]$, and otherwise delegates to an agent who plays the ex ante optimal action. We have shown that the candidate rule selects the optimal among the learning agents; it is left to verify that such a criterion for choosing between learning and non-learning agents is optimal for the principal.

Consider $\mu_p \geq \frac{1}{2}$; then among the non-learning agents, the principal would obviously choose the one who plays $a = R$ (rather than $a = L$), and such a choice yields the principal expected payoff $\mu_p \cdot 1 + (1-\mu_p)\cdot 0 = \mu_p$. Optimal delegation to a learning agent yields (by plugging the candidate rule into the principal's payoff obtained above)

$$\frac{e^{\frac{1}{\lambda}}}{e^{\frac{2}{\lambda}}-1}\left[e^{\frac{1}{\lambda}} - \mu_p\frac{1-\mu^*}{\mu^*} - (1-\mu_p)\frac{\mu^*}{1-\mu^*}\right] = \frac{e^{\frac{1}{\lambda}}}{e^{\frac{2}{\lambda}}-1}\left[e^{\frac{1}{\lambda}} - 2\sqrt{\mu_p(1-\mu_p)}\right]. \tag{2.28}$$

Taking the difference between (2.28) and μ_p, the payoff from delegating to a non-learning agent, let us find belief μ_p of a principal who would be indifferent between the two:

$$\begin{aligned}
&\frac{e^{\frac{1}{\lambda}}}{e^{\frac{2}{\lambda}}-1}\left[e^{\frac{1}{\lambda}} - 2\sqrt{\mu_p(1-\mu_p)}\right] - \mu_p = 0\\
\iff\ & e^{\frac{2}{\lambda}} - 2e^{\frac{1}{\lambda}}\sqrt{\mu_p(1-\mu_p)} = \mu_p e^{\frac{2}{\lambda}} - \mu_p\\
\iff\ & \left(e^{\frac{1}{\lambda}}\sqrt{1-\mu_p} - \sqrt{\mu_p}\right)^2 = 0\\
\iff\ & \frac{\sqrt{\mu_p}}{\sqrt{1-\mu_p}} = e^{\frac{1}{\lambda}}.
\end{aligned}$$

Hence, the principal prefers a learning agent when $\frac{\sqrt{\mu_p}}{\sqrt{1-\mu_p}} < e^{\frac{1}{\lambda}}$ and a non-learning agent when $\frac{\sqrt{\mu_p}}{\sqrt{1-\mu_p}} > e^{\frac{1}{\lambda}}$. Therefore, the candidate rule is indeed optimal for $\mu_p \geq \frac{1}{2}$. A mirror argument can be used to establish optimality for $\mu_p \leq \frac{1}{2}$. This concludes the proof of Proposition 2.2.

2.A.3 Proof of Lemma 2.1

The goal is to find the optimal choice probabilities $\beta^* \in \Delta(\mathcal{A})$ which maximize the principal's expected utility (2.17). First, let us rewrite expression (2.17) using $\delta \equiv e^{\frac{1}{\lambda}} - 1$:

$$\begin{aligned}\sum_{j=1}^{N} \mu_p(\omega_j) \frac{\beta(a_j) e^{\frac{1}{\lambda}}}{1+\delta\beta(a_j)} &= \sum_{j\in C(\beta)} e^{\frac{1}{\lambda}} \frac{\frac{\mu_p(\omega_j)}{\delta}(1+\delta\beta(a_j)) - \frac{\mu_p(\omega_j)}{\delta}}{1+\delta\beta(a_j)} \\ &= \sum_{j\in C(\beta)} e^{\frac{1}{\lambda}} \left(\frac{\mu_p(\omega_j)}{\delta} - \frac{\mu_p(\omega_j)}{\delta\left(1+\delta\beta(a_j)\right)} \right).\end{aligned}$$

The first term in the brackets above is independent of β, so the principal's maximization problem is equivalent to

$$\min_{\beta} \sum_{j\in C(\beta)} \frac{\mu_p(\omega_j)}{1+\delta\beta(a_j)}. \tag{2.29}$$

Let ξ denote the Lagrange multiplier corresponding to the constraint $\sum\limits_{j=1}^{N} \beta(a_j) = 1$. Then the first-order condition for $\beta(a_i)$ with $i \in C(\beta)$ is

$$(1+\delta\beta(a_i))^2 = -\frac{\mu_p(\omega_i)}{\xi}. \tag{2.30}$$

Summing up these equalities over all $j \in C(\beta)$, we get that

$$\sum_{j\in C(\beta)} (1+\delta\beta(a_j))^2 = -\frac{\sum_{j\in C(\beta)} \mu_p(\omega_j)}{\xi} \tag{2.31}$$

Combining (2.30) and (2.31):

$$1+\delta\beta(a_i) = \frac{\sqrt{\mu_p(\omega_i)}}{\sqrt{\sum_{j\in C(\beta)} \mu_p(\omega_j)}} \sqrt{\sum_{j\in C(\beta)} (1+\delta\beta(a_j))^2} \tag{2.32}$$

Once again summing up these equalities over all $j \in C(\beta)$, we get that

$$K(\beta) + \delta = \frac{\sum_{j\in C(\beta)} \sqrt{\mu_p(\omega_j)}}{\sqrt{\sum_{j\in C(\beta)} \mu_p(\omega_j)}} \sqrt{\sum_{j\in C(\beta)} (1 + \delta\beta(a_j))^2}.$$

Expressing $\sqrt{\sum_{j\in C(\beta)}(1 + \delta\beta(a_j))^2}$ from this expression and plugging it into (2.32) allows us to express $\beta(a_i)$ (for $i \in C(\beta)$) in closed form as

$$\beta(a_i) = \frac{1}{\delta}\left(\frac{(K(\beta) + \delta)\sqrt{\mu_p(\omega_i)}}{\sum_{j\in C(\beta)} \sqrt{\mu_p(\omega_j)}} - 1\right). \tag{2.33}$$

The necessary condition for option i to be in a consideration set ($i \in C(\beta)$) is $\beta(a_i) \geq 0$ or, equivalently,

$$\sqrt{\mu_p(\omega_i)} > \frac{1}{K(\beta) + \delta} \sum_{j\in C(\beta)} \sqrt{\mu_p(\omega_j)}.$$

Now let ξ_k denote the Lagrange multiplier for the constraint $\beta(a_k) \geq 0$. Then the first-order condition for an alternative $k \notin C(\beta)$ that is not chosen is

$$\mu_p(\omega_k) = -\xi - \xi_k \qquad \Rightarrow \qquad \mu_p(\omega_k) \leq -\xi.$$

Plugging in ξ from (2.30) into the inequality above yields

$$\mu_p(\omega_k) \leq \frac{\sum_{j\in C(\beta)} \mu_p(\omega_j)}{\sum_{j\in C(\beta)}(1 + \delta\beta(a_j))^2} \qquad \Leftrightarrow \qquad \sqrt{\mu_p(\omega_k)} \leq \frac{1}{K(\beta) + \delta} \sum_{j\in C(\beta)} \sqrt{\mu_p(\omega_j)}$$

for all $k \notin C(\beta)$.

Since the minimization problem has a convex objective function and linear constraints, the Kuhn-Tucker conditions are necessary and sufficient. Thus the necessary and sufficient conditions that the solution β^* must satisfy are given by:

$$\begin{cases} \sqrt{\mu_p(\omega_i)} > \frac{1}{K(\beta^*)+\delta} \sum\limits_{j\in C(\beta^*)} \sqrt{\mu_p(\omega_j)} & \text{for all } i \in C(\beta^*), \\ \sqrt{\mu_p(\omega_k)} \leq \frac{1}{K(\beta^*)+\delta} \sum\limits_{j\in C(\beta^*)} \sqrt{\mu_p(\omega_j)} & \text{for all } k \notin C(\beta^*). \end{cases}$$

Recall that we assumed, without loss of generality, that $\mu_p(\omega_1) \geq \mu_p(\omega_2) \geq \ldots \geq$

$\mu_p(\omega_N)$. Suppose that the solution β^* is such that $K(\beta^*) = K'$. Clearly then, in the optimum, the consideration set $C(\beta^*)$ will consist of the first K' alternatives.

Denote $\Delta_L \equiv (L+\delta)\sqrt{\mu_p(\omega_L)} - \sum_{j=1}^{L} \sqrt{\mu_p(\omega_j)}$. Notice that for all $L > 1$:

$$\begin{aligned}
\Delta_L \equiv &(L+\delta)\sqrt{\mu_p(\omega_L)} - \sum_{j=1}^{L} \sqrt{\mu_p(\omega_j)} \\
= &(L-1+\delta)\sqrt{\mu_p(\omega_{L-1})} - \sum_{j=1}^{L-1} \sqrt{\mu_p(\omega_j)} - \sqrt{\mu_p(\omega_L)} \\
&- (L-1+\delta)\sqrt{\mu_p(\omega_{L-1})} + (L+\delta)\sqrt{\mu_p(\omega_L)} \\
= &\Delta_{L-1} - (L-1+\delta)\left(\sqrt{\mu_p(\omega_{L-1})} - \sqrt{\mu_p(\omega_L)}\right).
\end{aligned}$$

Therefore, Δ_L decreases in L. Since $\Delta_1 > 0$, there either exists unique K' such that $\Delta_{K'} > 0$ and $\Delta_{K'+1} \leq 0$, or $\Delta_L > 0$ for all L. In the former case, $K(\beta^*) = K'$, and in the latter case, $K(\beta^*) = N$.

In the end, the solution to the principal's problem is given by $\beta^*(a_i)$ as in (2.33) if $i \in C(\beta^*)$, $\beta^*(a_i) = 0$ if $i \notin C(\beta^*)$, and $C(\beta^*) = 1, ..., K(\beta^*)$, where $K(\beta^*)$ is as described above.

2.A.4 Proof of Theorem 2.1

Corollary 2 from Matějka and McKay (2015) shows that a vector of the unconditional choice probabilities $\beta \in \Delta(\mathcal{A})$ solves (2.14) only if it solves the system of equations given by

$$\sum_{j=1}^{N} \mu(\omega_j) \frac{e^{\frac{u(a_i,\omega_j)}{\lambda}}}{\sum_{k=1}^{N} \beta(a_k) e^{\frac{u(a_k,\omega_j)}{\lambda}}} = 1, \tag{2.34}$$

for every $i \in \{1, ..., N\}$ such that $\beta(a_i) > 0$.

The question then is: given a vector $\beta \in \Delta(\mathcal{A})$ of unconditional choice probabilities, can we find $\mu \in \mathbb{R}_+^N$ that solves the following system:

$$\begin{cases} \mu(\omega_1) + \mu(\omega_2) + \ldots + \mu(\omega_N) = 1, \\ \sum\limits_{j=1}^{N} \mu(\omega_j) \frac{e^{\frac{u(a_i,\omega_j)}{\lambda}}}{\sum\limits_{k=1}^{N} \beta(a_k) e^{\frac{u(a_k,\omega_j)}{\lambda}}} = 1 & \forall i \in C(\beta). \end{cases} \tag{2.35}$$

The system above is a linear system of $K(\beta)+1$ equations with N unknowns. To prove the solution exists, we use the Farkas' lemma (see, e.g., Corollary 5.85 in Aliprantis and Border (2006)). It states that given some matrix $A \in \mathbb{R}^{m\times n}$ and a vector $b \in \mathbb{R}^m$, the linear system $Ax = b$ has a non-negative root $x \in \mathbb{R}^n_+$ if and only if there exists no vector $y \in \mathbb{R}^m$ such that $A'y \geq 0$ with $b'y < 0$. The two latter inequalities applied to our case form the following system:

$$\begin{cases} y_0 \left(\sum\limits_{k=1}^{N} \beta(a_k) e^{\frac{u(a_k,\omega_j)}{\lambda}} \right) + \left(\sum\limits_{i\in C(\beta)} y_i e^{\frac{u(a_i,\omega_j)}{\lambda}} \right) \geq 0 \quad \forall j \in \{1,...,N\}, \\ y_0 + \sum\limits_{i\in C(\beta)} y_i < 0. \end{cases} \tag{2.36}$$

We need to show there exists no $y \in \mathbb{R}^{K(\beta)+1}$ that solves the system above. Let us define $z_i \equiv y_i + y_0\beta(a_i)$ for $i \in C(\beta)$. Then, recalling that $e^{\frac{u(a_i,\omega_i)}{\lambda}} = e^{\frac{1}{\lambda}}$ and $e^{\frac{u(a_i,\omega_j)}{\lambda}} = 1$ for $i \neq j$, system (2.36) transforms to

$$\begin{cases} z_j e^{\frac{1}{\lambda}} + \sum\limits_{i\in C(\beta)\backslash\{j\}} z_i \geq 0 & \forall j \in C(\beta), \\ \sum\limits_{i\in C(\beta)} z_i \geq 0 & \forall j \in \{1,...,N\}\backslash C(\beta), \\ \sum\limits_{i\in C(\beta)} z_i < 0. \end{cases} \tag{2.37}$$

System (2.37) above does not have a solution. Indeed, if $C(\beta) \subsetneq \{1,...,N\}$, then the middle set of inequalities directly contradicts the latter inequality. If $C(\beta) = \{1,...,N\}$, then subtracting the latter inequality from the former, for a given $j \in C(\beta)$, yields $z_j\delta \geq 0 \iff z_j \geq 0$. Since this must hold for all $j \in C(\beta)$, we obtain a contradiction with the latter inequality, $\sum\limits_{i\in C(\beta)} z_i < 0$.

By the Farkas' lemma, we then conclude that for any vector $\beta \in \Delta(\mathcal{A})$ there exists a belief $\mu \in \mathit{\Delta}(\Omega)$ that solves system (2.35). This concludes the proof.

2.A.5 Proof of Proposition 2.3

This proof proceeds in two parts. First, we show that the delegation strategy introduced in the proposition (hereinafter referred to as "the candidate strategy") is optimal for the principal. Then we establish that it does indeed possess the stated properties.

Consider an agent with a prior belief

$$\mu(\omega_i) = \frac{\sqrt{\mu_p(\omega_i)}}{\sum_{j=1}^{N} \sqrt{\mu_p(\omega_j)}}. \tag{2.38}$$

It is trivial to verify that prior belief μ defined this way satisfies the candidate strategy in the statement of the proposition, and hence represents the candidate strategy. Consider an agent hired in accordance with the candidate rule. Substituting (2.38) into (2.15) yields

$$\beta(a_i) = \max\left\{0, \frac{1}{\delta}\left(\frac{(K(\beta^*)+\delta)\sqrt{\mu_p(\omega_i)}}{\sum_{j\in C(\beta^*)} \sqrt{\mu_p(\omega_j)}} - 1\right)\right\}, \tag{2.39}$$

which are exactly the probabilities stated in Lemma 2.1. Therefore, an agent hired according to the candidate strategy makes decisions in such a way that generates the principal-optimal unconditional choice probabilities. Therefore, delegation according to the candidate strategy is indeed optimal for the principal.

Now we show that the candidate strategy satisfies the properties stated in the proposition. Firstly, it follows clearly from (2.38) that $\mu^*(\omega_1) \geq \mu^*(\omega_2) \geq \ldots \geq \mu^*(\omega_N)$. It remains to show that $\mu^*(\omega_1) \leq \mu_p(\omega_1)$ and $\mu^*(\omega_N) \geq \mu_p(\omega_N)$. The former inequality can be shown as follows:

$$\begin{aligned}
&\mu^*(\omega_1) \leq \mu_p(\omega_1) \\
\iff &\frac{\sqrt{\mu_p(\omega_i)}}{\sum_{j=1}^{N} \sqrt{\mu_p(\omega_j)}} \leq \mu_p(\omega_1) \\
\iff &1 \leq \sqrt{\mu_p(\omega_1)} \cdot \left(\sum_{j=1}^{N} \sqrt{\mu_p(\omega_j)}\right) \\
\iff &1 \leq \mu_p(\omega_1) + \sqrt{\mu_p(\omega_1)\mu_p(\omega_2)} + \ldots + \sqrt{\mu_p(\omega_1)\mu_p(\omega_N)},
\end{aligned}$$

and the latter inequality holds because $\mu_p(\omega_1) + \ldots + \mu_p(\omega_N) = 1$ and $\sqrt{\mu_p(\omega_1)\mu_p(\omega_j)} \geq \mu_p(\omega_j)$ for all $j \in \{1, \ldots, N\}$, since $\mu_p(\omega_1) \geq \mu_p(\omega_j)$. Note that $\mu^*(\omega_1) = \mu_p(\omega_1)$ only if $\mu_p(\omega_1) = \ldots = \mu_p(\omega_N)$.

Similarly, the inequality $\mu^*(\omega_N) \geq \mu_p(\omega_N)$ is equivalent to

$$1 \geq \sqrt{\mu_p(\omega_1)\mu_p(\omega_N)} + ... + \sqrt{\mu_p(\omega_{N-1})\mu_p(\omega_N)} + \mu_p(\omega_N),$$

which holds because $\sqrt{\mu_p(\omega_j)\mu_p(\omega_N)} \leq \mu_p(\omega_j)$ for all $j \in \{1, ..., N\}$, with equalities only if $\mu_p(\omega_1) = ... = \mu_p(\omega_N)$. This concludes the proof of Proposition 2.3.

2.A.6 Proof of Proposition 2.4

It follows from (2.19) that the size of the consideration set in the aligned problem, $\bar{K}$, is such that

$$\sum_{j=1}^{\bar{K}} \frac{\mu_p(\omega_j)}{\mu_p(\omega_{\bar{K}})} < \bar{K} + \delta \leq \sum_{j=1}^{\bar{K}} \frac{\mu_p(\omega_j)}{\mu_p(\omega_{\bar{K}+1})}$$

Since $\frac{\mu_p(\omega_i)}{\mu_p(\omega_{\bar{K}})} > 1$ for all $i < \bar{K}$, we have that $\frac{\mu_p(\omega_i)}{\mu_p(\omega_{\bar{K}})} > \frac{\sqrt{\mu_p(\omega_i)}}{\sqrt{\mu_p(\omega_{\bar{K}})}} > 1$ holds for all $i < K$. Therefore,

$$\sum_{j=1}^{\bar{K}} \frac{\sqrt{\mu_p(\omega_j)}}{\sqrt{\mu_p(\omega_{\bar{K}})}} < \bar{K} + \delta. \tag{2.40}$$

From (2.20), K^* is the unique solution of

$$\sum_{j=1}^{K^*} \frac{\sqrt{\mu_p(\omega_j)}}{\sqrt{\mu_p(\omega_{K^*})}} < K^* + \delta \leq \sum_{j=1}^{K^*} \frac{\sqrt{\mu_p(\omega_j)}}{\sqrt{\mu_p(\omega_{K^*+1})}}. \tag{2.41}$$

Two cases are possible, depending on whether

$$\bar{K} + \delta \gtreqless \sum_{j=1}^{\bar{K}} \frac{\sqrt{\mu_p(\omega_j)}}{\sqrt{\mu_p(\omega_{\bar{K}+1})}}. \tag{2.42}$$

If $\bar{K} + \delta \leq RHS$ in (2.42) (where RHS refers to the right-hand side), then together with (2.40) this implies that $\bar{K}$ solves (2.41), and thus $\bar{K} = K^*$, which satisfies that statement of the proposition.

If, however, $\bar{K} + \delta > RHS$ in (2.42), then $\bar{K}$ does not solve (2.41). In this case, note that going from K by $K+1$ increases the LHS of (2.42) by 1 and increases the RHS by the amount strictly greater than 1, since a new term $\frac{\sqrt{\mu_p(\omega_{K+1})}}{\sqrt{\mu_p(\omega_{K+2})}} > 1$ is added to the sum, and all existing terms increase because $\mu_p(\omega_{K+1}) < \mu_p(\omega_K)$. This holds for all K, meaning that if $\bar{K} + \delta > RHS$ in (2.42), then $K + \delta > \sum_{j=1}^{K} \frac{\sqrt{\mu_p(\omega_j)}}{\sqrt{\mu_p(\omega_{K+1})}}$ for all $K < \bar{K}$. Therefore, the unique solution K^* of (2.41) must be such that $K_M > \bar{K}$. This concludes the proof.

2.A.7 Proof of Proposition 2.5

Part 2 of the statement follows immediately from Proposition 3 of Matveenko and Mikhalishchev (2021).

To show part 1, we invoke Theorem 1 from Matějka and McKay (2015) stated in (2.14), which claims that in the contracting problem, the $\beta : \Omega \to \Delta(\mathcal{A})$ that solves the agent's problem (2.21) is given by

$$\begin{aligned} \pi(a_i|\omega_j) &= \frac{\beta(a_i)e^{\frac{u(a_i,\omega_j)+\tau(a_i)}{\lambda}}}{\sum_{k=1}^{N} \beta(a_k)e^{\frac{u(a_k,\omega_j)+\tau(a_k)}{\lambda}}} \\ &= \frac{\beta'(a_i)e^{\frac{u(a_i,\omega_j)}{\lambda}}}{\sum_{k=1}^{N} \beta'(a_k)e^{\frac{u(a_k,\omega_j)}{\lambda}}}, \\ \text{where } \beta(a_i) &= \sum_{j=1}^{N} \mu(\omega_j)\pi(a_i|\omega_j). \\ \text{and } \beta'(a_i) &\equiv \frac{\beta(a_i)e^{\frac{\tau(a_i)}{\lambda}}}{\sum_{k=1}^{N} \beta(a_k)e^{\frac{\tau(a_k)}{\lambda}}}. \end{aligned} \tag{2.43}$$

Since β' is a valid probability distribution on $\mathcal{A}$, representation (2.43) together with (2.14) imply that such a collection of conditional probabilities π is a valid solution to the agent's problem (2.4) when the agent's preferences net of information costs are given by $u(a_i, \omega_j)$. That is, the principal can implement the desired conditional choice probabilities π by choosing an agent with unbiased preferences and some belief μ, such that the unconditional choice probabilities selected by this agent are given by β'. Theorem 2.1 implies that such a belief $\mu \in \Delta(\Omega)$ does indeed exist.

2.A.8 Proof of Proposition 2.6

Plugging (2.13) in (2.11) yields the optimal conditional choice probabilities for the binary model, given by

$$\begin{aligned}\pi^*(R|r) &= \left(e^{\frac{2}{\lambda}}-1\right)^{-1} e^{\frac{1}{\lambda}}\left(e^{\frac{1}{\lambda}}-\sqrt{\frac{1-\mu_p}{\mu_p}}\right),\\ \pi^*(L|l) &= \left(e^{\frac{2}{\lambda}}-1\right)^{-1} e^{\frac{1}{\lambda}}\left(e^{\frac{1}{\lambda}}-\sqrt{\frac{\mu_p}{1-\mu_p}}\right),\end{aligned} \tag{2.44}$$

cropped to $[0,1]$.

The agent's preferences only depend on the difference $\tau(R)-\tau(L)$. Assuming all $\tau(R)\in\mathbb{R}$ are available to the principal (no limited liability), it is without loss to set $\tau(L)=0$. The agent's problem is given by (2.21). Solving it given $\tau=(\tau(R),0)$ yields

$$\begin{aligned}\pi(R|r) &= 1-\frac{e^{\frac{2}{\lambda}}(1-\mu)-e^{\frac{1+\tau(R)}{\lambda}}+\mu}{\left(e^{\frac{2}{\lambda}}-1\right)\left(e^{\frac{1+\tau(R)}{\lambda}}-1\right)\mu},\\ \pi(L|l) &= \frac{e^{\frac{1}{\lambda}}\left(e^{\frac{2}{\lambda}}(1-\mu)-e^{\frac{1+\tau(R)}{\lambda}}+\mu\right)}{\left(e^{\frac{2}{\lambda}}-1\right)\left(e^{\frac{1}{\lambda}}-e^{\frac{\tau(R)}{\lambda}}\right)(1-\mu)},\end{aligned} \tag{2.45}$$

cropped to $[0,1]$.

The principal's contracting problem (2.22) in the binary setting with $\rho=0$ is similar to (2.12):

$$\begin{aligned}&\max_{\tau(R)}\left\{\mu_p\pi(R|r)+(1-\mu_p)\,\pi(L|l)\right\}\\ &\text{s.t. } \pi(R|r),\pi(L|l) \text{ are given by (2.45).}\end{aligned} \tag{2.46}$$

Assuming the probabilities in (2.45) are interior, the F.O.C. for (2.46) yields the candidate solution $\tau(R)$ given by

$$\tau^*(R)=\lambda\ln\left[\frac{\frac{1-\mu}{\mu}e^{\frac{1}{\lambda}}+\sqrt{\frac{1-\mu_p}{\mu_p}}}{\frac{1-\mu}{\mu}+e^{\frac{1}{\lambda}}\sqrt{\frac{1-\mu_p}{\mu_p}}}\right], \tag{2.47}$$

where the expression under the $\ln(\cdot)$ is non-negative for any λ,μ_p,μ, and thus the candidate $\tau(R)$ exists for any μ that yields interior probabilities (2.45).

Plugging (2.47) into (2.45) yields, after some routine manipulations, the conditional choice probabilities that coincide with (2.44) (hence, the probabilities (2.45) are interior given μ and $\tau^*(R)$ if and only if the probabilities (2.44) are interior). Thus, the condition

(2.47) is not only necessary, but also sufficient. Hence, for any μ_p for which (2.44) are interior, $\tau^*(R)$ as given by (2.47) solves the principal's problem (2.46), and this solution exists for any μ.

If λ and μ_p are such that probabilities (2.44) are not interior, then the principal would like the agent to take the ex ante (principal-)preferred action (it can be verified that the expressions in (2.44) are such that $\pi^*(R|r) \geq 1 \iff \pi^*(L|l) \leq 0$ and vice versa). The candidate transfers (2.47) yield exactly such non-interior probabilities (when plugged into (2.45)), and hence they still solve the principal's problem (2.46) for any respective μ.[30] This concludes the proof of part 1 of the proposition.

To show part 2, consider (2.47) as a function of μ. It is strictly decreasing in μ on $[0,1]$, and the equation $\tau^*(R)(\mu) = 0$ has a unique root in $[0,1]$ equal to

$$\mu^* = \frac{\sqrt{\mu_p}}{\sqrt{\mu_p} + \sqrt{1-\mu_p}},$$

meaning that $\tau(R) \geq 0 \iff \mu \leq \mu^*$.

2.A.9 Proof of Proposition 2.7

Proceeding analogously to Section 2.4, we obtain that the agent's problem (2.23) given incentive payment $\bar{\tau}$ (assuming $\underline{\tau} = 0$) is solved by

$$\begin{aligned} \pi\left(R|r\right) &= \frac{e^{\frac{1+\bar{\tau}}{\lambda}}\left(e^{\frac{1+\bar{\tau}}{\lambda}}\mu - (1-\mu)\right)}{\left(e^{\frac{2(1+\bar{\tau})}{\lambda}} - 1\right)\mu} = \frac{e^{\frac{1+\bar{\tau}}{\lambda}}}{e^{\frac{2(1+\bar{\tau})}{\lambda}} - 1}\left(e^{\frac{1+\bar{\tau}}{\lambda}} - \frac{1-\mu}{\mu}\right), \\ \pi\left(L|l\right) &= \frac{e^{\frac{1+\bar{\tau}}{\lambda}}\left(e^{\frac{1+\bar{\tau}}{\lambda}}(1-\mu) - \mu\right)}{\left(e^{\frac{2(1+\bar{\tau})}{\lambda}} - 1\right)(1-\mu)} = \frac{e^{\frac{1+\bar{\tau}}{\lambda}}}{e^{\frac{2(1+\bar{\tau})}{\lambda}} - 1}\left(e^{\frac{1+\bar{\tau}}{\lambda}} - \frac{\mu}{1-\mu}\right), \end{aligned} \tag{2.48}$$

each cropped to $[0,1]$.

The principal's problem (2.24) can be rewritten as

$$\begin{aligned} &\max_{\bar{\tau}} \left\{(1-\bar{\tau})\left(\mu_p \pi(R|r) + (1-\mu_p)\pi(L|l)\right)\right\}, \\ &\text{s.t. } \pi(R|r), \pi(L|l) \text{ are given by (2.48).} \end{aligned} \tag{2.49}$$

[30] Note that $\tau^*(R)$ is not the unique solution in this case. If $\pi^*(R|r) = 1, \pi^*(L|l) = 0$, then any $\tau(R) \geq \lambda \ln\left(\mu + (1-\mu)e^{\frac{2}{\lambda}}\right) - 1$ yields the optimal choice probabilities, and if $\pi^*(R|r) = 0, \pi^*(L|l) = 1$, then any $\tau(R) \leq 1 - \lambda \ln\left(\mu e^{\frac{2}{\lambda}} + (1-\mu)\right)$ solves the principal's problem.

To begin with, note that the principal would never choose $\bar{\tau} < 0$ (due to limited liability) or $\bar{\tau} \geq 1$ (since this would render the principal's payoff negative). Further, if the solution (2.48) yields values outside $[0, 1]$ for some $\bar{\tau} > 0$, then such a $\bar{\tau}$ is clearly suboptimal for the principal, as she can then reduce her costs without decreasing the precision of the agent's choice. Thus, in the optimum, if $\bar{\tau} > 0$, then the probabilities (2.48) are interior. Assuming the latter and plugging (2.48) into the principal's problem (2.49), the F.O.C. of this problem is given by

$$\mu_p \frac{1-\mu}{\mu} + (1-\mu_p)\frac{\mu}{1-\mu} = e^{\frac{1+\bar{\tau}}{\lambda}} \cdot \frac{\lambda\left(e^{2\frac{1+\bar{\tau}}{\lambda}} - 1\right) + 2(1-\bar{\tau})}{\lambda\left(e^{2\frac{1+\bar{\tau}}{\lambda}} - 1\right) + \left(e^{2\frac{1+\bar{\tau}}{\lambda}} + 1\right)(1-\bar{\tau})}. \tag{2.50}$$

Let $\gamma(\mu, \mu_p)$ denote the LHS and $\chi(\bar{\tau}, \lambda)$ the RHS of (2.50), respectively. Then the necessary condition for an interior $\bar{\tau}$ to be optimal is $\gamma(\mu, \mu_p) = \chi(\bar{\tau}, \lambda)$. Note that $\gamma(\mu, \mu_p)$ is minimized for a given μ_p by $\mu = \mu^*(\mu_p)$ given by (2.13), and $\gamma(\mu^*(\mu_p), \mu_p) = 2\sqrt{\mu_p(1-\mu_p)} < 1$. Further, it can be shown that $\chi(\bar{\tau}, \lambda)$ is continuous and increasing in $\bar{\tau}$ for all λ. This means that a solution to (2.50) exists for a given μ, μ_p, λ if and only if $\chi(0, \lambda) \leq \gamma(\mu, \mu_p) \leq \chi(1, \lambda)$ (where the "only if" part follows from the intermediate value theorem). Since $\chi(\bar{\tau}, \lambda)$ is increasing in $\bar{\tau}$, the second-order condition holds, meaning that any $\bar{\tau}$ that solves (2.50) is a local maximizer of (2.49); further, there is at most one local maximizer.

Suppose $\mu < \mu^*(\mu_p)$. As $1 \leq \chi(0, \lambda) < +\infty$ for all $\lambda > 0$, $\gamma(0, \mu_p) = +\infty$, $\gamma(\mu^*(\mu_p), \mu_p) < 1$, and $\gamma(\mu, \mu_p)$ is continuous in $\mu \in (0, 1)$, there exists $\bar{\mu}_1 < \mu^*(\mu_p)$ such that $\gamma(\bar{\mu}_1, \mu_p) = \chi(0, \lambda)$. As $\chi(1, \lambda) = e^{\frac{2}{\lambda}} > 1$, there also exists $\tilde{\mu}_1 < \mu^*(\mu_p)$ such that $\gamma(\tilde{\mu}_1, \mu_p) = \chi(1, \lambda)$. As $\chi(0, \mu) < \chi(1, \lambda)$ for all λ, it follows that $\tilde{\mu}_1 < \bar{\mu}_1$. By a mirror argument, there also exist $\tilde{\mu}_2 > \bar{\mu}_2 > \mu^*(\mu_p)$ that satisfy the analogous properties. Further, $\gamma(\mu_p, \mu_p) = 1$, so that $\bar{\mu}_2 > \mu_p$. In the end, $\chi(0, \lambda) \leq \gamma(\mu, \mu_p) \leq \chi(1, \lambda)$ and, hence, a solution $\bar{\tau}$ to (2.50) exists (and thus constitutes a local maximum of (2.49)) if and only if

$$\mu \in [\tilde{\mu}_1, \bar{\mu}_1] \cup [\bar{\mu}_2, \tilde{\mu}_2]. \tag{2.51}$$

If μ does not satisfy (2.51) (which is the case in the "aligned" case, $\mu = \mu_p$), no interior solution exists to (2.49), hence $\bar{\tau} = 0$ is optimal. On the other hand, if μ satisfies (2.51), then the optimal $\bar{\tau}$ (that solves (2.49)) can be given by either the solution to (2.50), or the corner solution $\bar{\tau} = 0$. As argued previously, the latter can only be a candidate solution if the probabilities (2.48) are non-interior given $\bar{\tau} = 1$, which is the case if $\mu \notin \left[\frac{1}{1+e^{\frac{1}{\lambda}}}, \frac{e^{\frac{1}{\lambda}}}{1+e^{\frac{1}{\lambda}}}\right]$. Further, it can be shown that $\tilde{\mu}_1 < \frac{1}{1+e^{\frac{1}{\lambda}}}$ and $\frac{e^{\frac{1}{\lambda}}}{1+e^{\frac{1}{\lambda}}} < \tilde{\mu}_2$. Thus, we conclude that if

$$\mu \in \left[\frac{1}{1+e^{\frac{1}{\lambda}}}, \bar{\mu}_1\right] \cup \left[\bar{\mu}_2, \frac{e^{\frac{1}{\lambda}}}{1+e^{\frac{1}{\lambda}}}\right], \tag{2.52}$$

then a $\bar{\tau} \in (0,1)$ that solves (2.50) is a global maximizer of (2.49), which proves the statement of the proposition. It is, however, worth noting that (2.52) is only a sufficient condition and not a necessary one. Further, one of the invervals in (2.52) may be empty for extreme enough μ_p.

2.A.10 Proof of Proposition 2.8

Using Theorem 1 of Caplin et al. (2019), the agent's problem (2.25) given some restriction set A^* is solved by π such that the corresponding $\beta \in \Delta(A^*)$ satisfies (2.18) for all $a_i \in A^*$. Further, recall from Section 2.5 that π and β are connected in the optimum by relation (2.14) (where we set $\pi(a_i|\omega_j) \equiv \beta(a_i) \equiv 0$ for all $a_i \notin A^*$ and all $\omega_j \in \Omega$). Then by plugging (2.14) and the state-matching utility into the principal's expected payoff, it can be rewritten as in (2.17):

$$\sum_{i=1}^{N} \mu_p(\omega_i) \frac{\beta(a_i) e^{\frac{1}{\lambda}}}{1+\delta\beta(a_i)} = \sum_{i \in C(\beta)} \mu_p(\omega_i) \frac{(1+\delta)\beta(a_i)}{1+\delta\beta(a_i)}.$$

Plugging in (2.18) for β in the expression above transforms it to

$$\sum_{i \in C(\beta)} \frac{\frac{1+\delta}{\delta}\mu_p(\omega_i)\left[(K(\beta)+\delta)\,\mu_p(\omega_i) - \sum_{j \in C(\beta)} \mu(\omega_j)\right]}{(K(\beta)+\delta)\mu_p(\omega_i)}$$
$$= \frac{1+\delta}{\delta}\left[\sum_{i \in C(\beta)} \mu_p(\omega_i) - \sum_{i \in C(\beta)} \frac{\sum_{j \in C(\beta)} \mu(\omega_j)}{(K(\beta)+\delta)}\right]$$
$$= \frac{1+\delta}{K(\beta)+\delta} \sum_{i \in C(\beta)} \mu_p(\omega_i). \tag{2.53}$$

To prove the proposition statement, we proceed by induction. Consider some arbitrary action set $A_- \subset \mathcal{A}$ such that $a_k \notin A_-$ for some $k \in \{1, ..., N\}$ and another action set

$A_+ \equiv A_- \cup \{a_k\}$. Let β_+ denote the unconditional choice probabilities corresponding to the solution of (2.25) given A_+, let $C_+ \equiv C(\beta_+)$ and $K_+ \equiv K(\beta_+)$, and define β_-, C_-, K_- analogously given A_-.

Our goal is to show that that selecting $A^* = A_+$ is weakly better for the principal than $A^* = A_-$. If $a_k \notin C_+$, then the payoffs in the two cases are equal, and the statement is trivially true. Otherwise, using (2.53) for the principal's expected payoff, the statement amounts to:

$$\begin{aligned}
& 0 \leq \left(\frac{1+\delta}{K_+ + \delta} \sum_{i \in C_+} \mu_p(\omega_i)\right) - \left(\frac{1+\delta}{K_- + \delta} \sum_{i \in C_-} \mu_p(\omega_i)\right) \\
\iff & 0 \leq \left((K_- + \delta) \sum_{i \in C_+} \mu_p(\omega_i)\right) - \left((K_+ + \delta) \sum_{i \in C_-} \mu_p(\omega_i)\right) \\
\iff & 0 \leq (K_- + \delta)\mu_p(\omega_k) - \left(\sum_{i \in C_-} \mu_p(\omega_i)\right).
\end{aligned} \tag{2.54}$$

Since $a_k \in C_+$ by assumption, $\beta_+(a_k) > 0$, which, from (2.18), implies that

$$\begin{aligned}
& 0 < \frac{(K(\bar{\beta}) + \delta)\mu(\omega_i)}{\sum_{j \in C(\bar{\beta})} \mu(\omega_j)} - 1 \\
\iff & 0 < (K_+ + \delta)\mu_p(\omega_k) - \left(\sum_{i \in C_+} \mu_p(\omega_i)\right) \\
\iff & 0 < (K_- + 1 + \delta)\mu_p(\omega_k) - \left(\sum_{i \in C_-} \mu_p(\omega_i) + \mu_p(\omega_k)\right) \\
\iff & 0 < (K_- + \delta)\mu_p(\omega_k) - \left(\sum_{i \in C_-} \mu_p(\omega_i)\right),
\end{aligned}$$

which immediately implies that (2.54) holds. Therefore, it is indeed better for the principal to choose A_+ over A_-. Since A_- was arbitrary, this proves by induction that allowing a larger action set is always weakly better for the principal, and hence proves the original proposition.

2.A.11 Proof of Proposition 2.9

We provide an example for $N = 3$. We use the same version of the Farkas' Lemma as in the proof of Theorem 2.1. To show that there is no prior belief that solves the system of the first-order conditions for the problem, it is sufficient to show that there is a solution to the following dual inequality system

$$\begin{cases} z_1 e^{\frac{u(a_1,\omega_1)}{\lambda}} + z_2 e^{\frac{u(a_1,\omega_2)}{\lambda}} + z_3 e^{\frac{u(a_1,\omega_3)}{\lambda}} \geq 0, \\ z_1 e^{\frac{u(a_2,\omega_1)}{\lambda}} + z_2 e^{\frac{u(a_2,\omega_2)}{\lambda}} + z_3 e^{\frac{u(a_2,\omega_3)}{\lambda}} \geq 0, \\ z_1 e^{\frac{u(a_3,\omega_1)}{\lambda}} + z_2 e^{\frac{u(a_3,\omega_2)}{\lambda}} + z_3 e^{\frac{u(a_3,\omega_3)}{\lambda}} \geq 0, \\ z_1 + z_2 + z_3 < 0. \end{cases} \tag{2.55}$$

Let us normalize $\lambda = 1$ and consider payoffs given by the following matrix:

$$\begin{pmatrix} u(a_1,\omega_1) & u(a_2,\omega_1) & u(a_3,\omega_1) \\ u(a_1,\omega_2) & u(a_2,\omega_2) & u(a_3,\omega_2) \\ u(a_1,\omega_3) & u(a_2,\omega_3) & u(a_3,\omega_3) \end{pmatrix} = \begin{pmatrix} \ln 3 & 0 & \ln(2+\varepsilon) \\ 0 & \ln 3 & \ln(2+\varepsilon) \\ 0 & 0 & \ln(2+\varepsilon) \end{pmatrix}$$

Notice that vector $(z_1, z_2, z_3) = (-1-\delta, -1-\delta, 2)$ for small enough $\delta, \varepsilon \geq 0$ solves system (2.55): the two latter inequalities hold trivially for all such z, and the two former inequalities hold if $\varepsilon \geq 3^{\frac{1+\delta}{2}} - 2$. Therefore, there exists no μ that solves system (2.35) given $\beta \in \Delta(\Theta)$.

2.A.12 Proof of Proposition 2.10

We first show that there exists an equilibrium in the communication game that replicates the deletation equilibrium: the optimal agent acquires the same information, makes a truthful action recommendation, and the principal follows the recommendation.

Suppose that under delegation, the optimally chosen agent follows a decision rule β^* that yields a consideration set $C(\beta^*) = \{1, ..., K^*\}$. By Lemma 2.1, we have that

$$\begin{aligned} \sqrt{\mu(\omega_{K^*})} &\geq \frac{1}{K^*+\delta} \sum_{i=1}^{K^*} \sqrt{\mu(\omega_i)} \\ \iff \delta \sqrt{\mu(\omega_{K^*})} &\geq \sum_{i=1}^{K^*-1} \left(\sqrt{\mu(\omega_i)} - \sqrt{\mu(\omega_{K^*})} \right) \end{aligned} \tag{2.56}$$

Suppose the agent reports truthfully. Given the state-matching payoffs, for the principal to follow recommendation $\tilde{a} = \tilde{a}_{K^*}$ whenever it is issued, it must hold that

$$\mu_p(\omega_{K^*}|\tilde{a}_{K^*}) = \max_i \mu_p(\omega_i|\tilde{a}_{K^*}), \tag{2.57}$$

where $\mu_p(\omega|\tilde{a})$ is the probability that the principal's posterior belief assigns to state ω after hearing recommendation $\tilde{a}$ from the agent. In equilibrium, the principal's posterior $\mu_p(\omega_{K^*}|\tilde{a}_{K^*})$ must satisfy Bayes' rule:

$$\begin{aligned}
\mu_p(\omega_{K^*}|\tilde{a}_{K^*}) &= \frac{\pi(a_{K^*}|\omega_{K^*})\mu_p(\omega_{K^*})}{\sum_{i=1}^N \mu_p(\omega_i)\pi(a_{K^*}|\omega_i)} \\
&= \frac{\beta(a_{K^*})e^{\frac{1}{\lambda}}}{\beta(a_1) + ... + \beta(a_{K^*-1}) + \beta(a_{K^*})e^{\frac{1}{\lambda}}} \cdot \frac{\mu_p(\omega_{K^*})}{\sum_{i=1}^N \mu_p(\omega_i)\pi(a_{K^*}|\omega_i)} \\
&= \frac{\beta(a_{K^*})e^{\frac{1}{\lambda}}}{1 + \delta\beta(a_{K^*})} \cdot \frac{\mu_p(\omega_{K^*})}{\sum_{i=1}^N \mu_p(\omega_i)\pi(a_{K^*}|\omega_i)} \\
&= \frac{\sum_{i=1}^{K^*} \sqrt{\mu_p(\omega_i)}}{K^* + \delta} \cdot \beta(a_{K^*})e^{\frac{1}{\lambda}} \cdot \frac{\sqrt{\mu(\omega_{K^*})}}{\sum_{i=1}^N \mu_p(\omega_i)\pi(a_{K^*}|\omega_i)},
\end{aligned}$$

Where the last line is obtained by plugging the expression for $\beta(a_{K^*})$ from Lemma 2.1 in the denominator of the preceding line. Similarly, we can calculate the probability that the principal's posterior assigns to any other state ω_j:

$$\mu_p(\omega_j|\tilde{a}_{K^*}) = \begin{cases} \frac{\sum_{i=1}^{K^*} \sqrt{\mu_p(\omega_i)}}{K^*+\delta} \cdot \beta(a_{K^*})e^{\frac{1}{\lambda}} \cdot \frac{\sqrt{\mu(\omega_j)}}{\sum_{i=1}^N \mu_p(\omega_i)\pi(a_{K^*}|\omega_i)} & \text{if } j < K^*, \\ 0 & \text{if } j > K^*. \end{cases}$$

For condition (2.57) to hold, it is then enough for

$$e^{\frac{1}{\lambda}}\sqrt{\mu(\omega_{K^*})} \geq \sqrt{\mu(\omega_1)} \qquad \Longleftrightarrow \qquad \delta\sqrt{\mu(\omega_{K^*})} \geq \sqrt{\mu(\omega_1)} - \sqrt{\mu(\omega_{K^*})}, \tag{2.58}$$

to be satisfied. Note, however, that it is strictly weaker than (2.56), since

$$\sqrt{\mu(\omega_1)} - \sqrt{\mu(\omega_{K^*})} < \sum_{i=1}^{K^*-1} \left(\sqrt{\mu(\omega_i)} - \sqrt{\mu(\omega_{K^*})}\right).$$

Therefore, we conclude that (2.58) holds, and thus it is optimal for the principal to choose action a_{K^*} when the agent with prior belief μ^* recommends it.

Following the same argument, we can show the same for any other recommendation

$\tilde{a}_i$ for $i \in C(\beta^*)$: the necessary and sufficient condition for the principal to find it optimal to follow the recommendation would be

$$e^{\frac{1}{\lambda}}\sqrt{\mu(\omega_i)} \geq \sqrt{\mu(\omega_1)},$$

which is implied by (2.57), since $\mu(\omega_i) \geq \mu(\omega_{K^*})$ for $i \in C(\beta^*)$. This concludes the proof.

Chapter 3

Is it Better to be First? Search with Endogenous Information Acquisition

3.1 Introduction

Search theory is the classical approach to modeling information acquisition, and has been widely used in several fields of economics. The main focus of economic analysis has been the impact of search frictions.[1] However, only a little attention has been paid to the order effect in search theory. Usually, it is assumed that an agent searches mostly mechanically, choosing only the optimal order of inspection and when to stop a search procedure. In this paper we study the implications of the order of inspection in the endogenous information acquisition problem.

The standard assumption in such search models is that binary technology resolves uncertainty. That is, an agent can only obtain full information about the quality of an alternative by paying the fixed costs. This assumption substantially simplifies the analysis of the search problem but makes an agent's strategy more passive. In this paper, we relax the binary technology assumption and allow our agent to choose the nature of information about an alternative herself. Therefore, our agent can engage in a more active search strategy: she can endogenously choose the amount and the structure of information at every stage, and the information may be history-dependent. In this case, the order of alternatives influences her choice through the new endogenous information channel.

[1] See, e.g., Kamenica (2017) for the discussion.

In this paper, we build a stylized sequential model of the optimal interview design. In our model, the manager must choose one candidate for a job. Before making a choice, the manager may obtain information about candidates. The manager interviews one candidate at a time, and during the interview, she obtains information only about that candidate. The manager obtains information about candidates sequentially: after interviewing one candidate, she interviews another. We focus on the order effect and assume identical ex ante candidates. The structure of an interview is endogenous: the manager may learn different information about different candidates.

The endogenous information assumption is realistic in our setting. Although interview timeslots for candidates are usually prespecified, the manager does not have to obtain the same information during an interview with different candidates. She can endogenously select questions during an interview and thereby learn different information about the candidates. Moreover, the manager may choose to learn differently even when the candidates are ex ante similar, and the chosen information may depend on the path of inspections. Intuitively, the learning strategy allows the manager to more often choose choose some candidates according to their interview order, even if all candidates are ex ante similar.[2]

Using the stylized model of sequential interview design, we study the discrimination that the order of inspection can cause. To isolate the order effect, we focus on the setup with ex ante identical candidates. We say that there is discrimination if the choice rule, the unconditional probability of choice, is not uniform. For example, we say that there is discrimination against the last candidate if the manager unconditionally chooses him the least often. Additionally, taking inspiration from the psychological literature, we relate discrimination to the amount of information that the manager obtains. We first discuss our results regarding discrimination and later discuss its connection to the information obtained.

The main insight from previous economics research is that it is better to be first. There is discrimination against the last alternatives in related search settings. For example, Armstrong et al. (2009) show this in the oligopolistic search model, and Gossner et al. (2021) in the endogenous acquisition framework. The issue of discrimination caused by has also been up brought to attention in a public discussion. There is anecdotal evidence

[2]The presence of such discrimination creates strategic incentives for a candidate if he has additional information about his quality. However, in this paper, we focus only on the order effect in the manager's problem, and our candidates are passive. We abstract from all strategic considerations the order effect may cause.

of interest in how order influences choice discrimination in public internet forums. For example, in the popular internet forum `reddit.com`, many people ask for advice about which interview timeslot from those available they should ideally take. Usually, there is no universal advice from the respondents on which position is the most vulnerable to discrimination. However, the leading responses support basic common sense and can be summarized as "Just come in and do a good interview", suggesting that the order does not cause discrimination in choice.[3]

Our results support both possibilities. Depending on the number of available candidates, our findings distinguish the presence of discrimination. We show that there is no discrimination in a two candidate case: the manager chooses candidates uniformly on aggregate. With three candidates, it is better to be first: there is discrimination against the last candidate.[4]

In the case of two candidates, there are several optimal learning strategies, and the manager can conduct the interviews differently. With three candidates, the learning strategy is unique. We interpret the optimal strategy as interviews with decreasing difficulties. The manager conducts easier interviews in the later stages. In the interview design, the manager optimally balances the informativeness of the interview and its difficulty. Intuitively, if the manager conducts too difficult interviews in the first stages, they will not be informative enough. The candidate will fail too often, and the discrimination results will not hold.

The optimal structure of the interview in the case of three alternatives has an additional interesting feature. If the manager receives a good signal about the quality of a candidate in an interview, she does not consume informative signals in the interviews with later candidates. This pattern resembles the satisficing behavior from Simon (1955). However, in our case, the choice threshold decreases over time. Our manager engages in "cherry-picking" behavior. The manager aims to get better candidates during the first interviews.

Additionally, we relate the discrimination to the amount of information that the manager obtains.[5] Psychologists have studied the relation between the order and information

[3] `https://www.reddit.com/r/ITCareerQuestions/comments/3nk7o4/which_interview_slot_is_best_to_take/`

[4] To make a comparison between two cases more transparent in the three candidate case, we mainly analyze the learning strategy in two periods: the second and the third.

[5] In the optimum the information structures that the manager chooses in different periods are not comparable in a Blackwell sense. We measure and compare information in different periods in the amounts that the manager obtains.

in search problems in detail. The presence of a serial-position effect has been widely documented: see, e.g., Feigenbaum and Simon (1962). Their main insight is that a person who inspects items sequentially recalls different pieces of information about the items in relation to the order in which she inspected them. People tend to remember the first and last observations most strongly. The primacy effect presents when a person remembers more first inspections, and the recency effect presents when she remembers the last. However, the relationships between the serial-position effects and discrimination in choice is rather unclear. The serial-position effect can lead to discrimination against the first or the last alternatives; see, e.g., Li and Epley (2009), and Newell and Shanks (2014).

In our analysis, we use the amount of information a manager *learns* about alternatives as a proxy for the information a person *recalls* about alternatives. Simply, if a person learns more about an alternative, it is plausible that she will remember more information about it. We say that the primacy (recency) effect is present if the manager chooses to learn more (less) in the first stages. Our results differ for two cases. In the case with two alternatives, we find that the serial-position effect may be present but does not lead to discrimination. The exact conditions for the primacy or recency effects are not identifiable: the manager obtains a fixed amount of information in two periods and can obtain any positive amount in the first period and the rest in the second. In the optimum, the manager uses information as a perfect substitute between periods. In the case with three alternatives, the primacy effect is present. Therefore, the amount of obtained information is correlated with discrimination in choice.

Our analysis allows for flexibility of the information the manager can acquire. We assume that the manager is rationally inattentive. She may choose any information she wants, and pays some costs for it. To make the problem tractable, we model the quality of a candidate as a simple binary random variable, which takes values 0 and 1 with low and high quality interpretation. As in the theory of discrete rational inattention (Caplin et al. (2022)), we model the information acquisition strategy as a conditional distribution over signal realizations. We use the posterior approach with backward induction and concavification technique to solve the problem.

For tractability, we assume a quadratic information cost function form. This cost function falls into the class of the posterior-separable cost functions (Caplin et al. (2022)) and has been used previously in the rational inattention literature, see, e.g., Wei (2021), and Lipnowski et al. (2022). Our results are sensitive to the quadratic cost specification. Technically, in the case with two candidates, the manager's value function in the first period is linear in the posterior realization in the interval around the prior belief. This finding generates non-uniqueness of the learning strategy in the first period. For example, the manager is indifferent between not learning at all and having posterior realizations on the boundaries of the linear region.[6]

We study several simple extensions of our model. We consider a setup in which the expected qualities of the candidates are close to each other but not identical. The optimal interview structure is the same: there is a non-unique (unique) optimal strategy in the case of two (three) alternatives. In the case of two alternatives, the manager chooses the ex ante better candidate more often. The result of the substitutability of the information persists. Perhaps surprisingly, the manager is indifferent between inspecting the better or worse candidate first. This is not true in the classical search theory model with binary information technology (Weitzman (1979)). Interestingly, with three alternatives, the manager can choose the second candidate more often than the third, even if he is a worse ex-ante. However, we show that it is not optimal if the manager also chooses the order of inspection. The manager prefers to interview the better candidate first and chooses him more often.

If the manager discounts the future information, she does not use information as a perfect substitute in the case with two candidates. A recency effect is present: she consumes informative signals only in the second period. However, the no-discrimination result holds: the manager chooses candidates uniformly on average. If discounting is not huge, the optimal structure of an interview in the case with three alternatives holds. Interestingly, in this case, discounting leads to a perfect correlation between the serial-position effect and discrimination. We show that depending on the value of the discount factor, the primacy or the recency effect is present. If the primacy (recency) effect is present, the manager discriminates against the third (the second) candidate. Intuitively, if the information becomes much more expensive in the second period, the manager conducts a too difficult interview and chooses the second candidate too rarely.

Endogenous information acquisition is a natural assumption in many search settings, and we contribute to this relatively small but growing literature. The closest paper to ours is Jain and Whitmeyer (2021). They consider a search model in which an agent can

[6] In the paper we focus only on the interior learning strategy. Because the derivative of the cost function is bounded on the boundaries, full learning, in general, may be optimal. We show that it is possible to eliminate the full learning case by restricting the set of parameters. In general, the full learning strategy can be seen as the result of the optimization with active constraint and therefore can create non-trivial intertemporal effects. We found this case to be analytically intractable.

also learn flexibly, paying quadratic costs. However, they focus on the strategic incentives in the oligopolistic setting and market outcomes. Liu et al. (2022) consider a search model in which an agent may reduce the noise of her item evaluation. They mainly focus on comparing their model with classical search results, and they do not allow flexibility of information. In Dogan and Hu (2022), similarly to us, an agent has a flexible learning technology, but they study an equilibrium, in which an agent faces a stationary search environment. The literature on experimentation and sequential learning is also related to our model, e.g., Fershtman and Pavan (2022), Gossner et al. (2021); however, their modeling assumptions are quite different from ours.

Our model is a variant of the dynamic discrete rational inattention problem. Steiner et al. (2017) solve this type of problem using stochastic choices, and Miao and Xing (2020) consider a posterior approach. There is a substantial difference in our model. In both Steiner et al. (2017) and Miao and Xing (2020), an agent can learn at every stage any information about the payoff-relevant state. However, in our setup, an agent in stage i can learn only about alternative i, which we find a more realistic assumption in the optimal interview design application. In our setting, posterior belief about the alternative enters into the next period 's rewards, which does not happen in Steiner et al. (2017) and Miao and Xing (2020). For this reason, their results do not apply in our setup.

The rest of the paper is organized as follows: we formulate the two-period benchmark model with ex ante identical candidates in Section 3.2. We analyze the model in Section 3.3. We introduce an additional alternative to this model in Section 3.4. In Section 3.5 we explore the generalization of the results of the main model and consider several extensions, including different expected qualities, discounting, and an alternative cost function. In Section 3.6 we conclude.

3.2 Two-Period Model with Identical Candidates

There is a decision maker (manager, she) and two alternatives (candidates, he) indexed by i. A candidate i has a binary type $\theta_i = \{0, 1\}$ and the manager's prior belief is $\mathrm{P}(\theta_i = 1) = \mu$. Types are independently distributed.

The manager must choose one candidate, there is no outside option. Before making a decision, she may learn about the candidates' types. Learning is costly and sequential. In stage i the manager inspects candidate i. We assume no discounting between periods and therefore the total costs in the problem for the manager is simply the sum of costs in each period i.[7]The manager's payoff in the problem equals the difference between the expected value of the chosen candidate and total costs.

Learning in stage i proceeds as follows. The manager chooses a Blackwell experiment, which generates information about θ_i. This experiment is a mapping from the state space $\{0, 1\}$ to a probability measure over some compact set of signal realizations. Each signal realization is associated with posterior belief distribution on $\{0, 1\}$ and an experiment induces a distribution over posterior beliefs. Because we consider a binary random variable, we identify a posterior belief distribution with a belief about $\theta_i = 1$. It is commonly known from the literature on rational inattention (e.g., Caplin and Dean (2013)), that instead of considering the set of Blackwell experiments, one can consider the set of distribution of posterior beliefs, in which the mean equals prior. In our analysis, we apply the posterior approach.

Formally, the manager chooses distribution p_i, which satisfies Bayesian plausibility condition:

$$p_i \in \Delta[0,1] : \quad \int\limits_{[0,1]} x dp_i(x) = \mu.$$

Below we refer to a posterior distribution that satisfies this condition as a feasible distribution. In an optimization problem, we always assume feasibility and sometimes, for brevity, omit this requirement. A choice of p_i is costly. We assume quadratic cost specification, that is, if k is a marginal cost of information then the cost of information in stage i is

$$C(p_i) = k \int\limits_{[0,1]} (x - \mu)^2 dp_i(x).$$

This cost specification falls into the class of the posterior-separable cost function, the most commonly used function in the rational inattention literature (Caplin et al. (2022)).[8]This cost specification is proportional to the variance of the posterior belief and possesses a standard property: the cost increases if x moves closer to the certainty from prior μ.

Because the manager's problem is sequential, it is convenient to introduce the payoffs sequentially. We fix a first-period learning strategy p_1. After the first period, the manager

[7]We allow the manager to not learn and to spend zero costs in stages i. For example, let the manager incur some positive cost in the first period and suppose that her evaluation of the candidate exceeds μ. If she incurs zero cost in the second period, we interpret this to mean that the manager knowingly chose a candidate in the first period and in the second period simply she consumes no informative signals about a candidate.

[8]This specification is for binary random variable θ_i. It can be easily extended for a random variable with more realizations, such as $C(p_i) = k \int\limits_{[0,1]} ||x - \mu|| dp_i(x)$ as e.g., in Lipnowski et al. (2022)

obtains a posterior belief x_1 about the quality of the first candidate. This value plays the role of the outside option in the second period.

We define the net utility function in the second period given x_1 and posterior belief about the second candidate x_2 as

$$u_2(x_1, x_2) = \max\{x_1, x_2\} - k(x_2 - \mu)^2.$$

The expected utility of the manager in the second period is therefore the expectation over posterior realizations x_2. Thus the manager's problem is choosing the best feasible posterior distribution:

$$\max_{p_2 \in \Delta[0,1]} \int_{[0,1]} u_2(x_1, x) dp_2(x). \tag{3.1}$$

We denote the maximum attained value of the problem as $\mathcal{V}(x_1)$. Similarly to the second period, we define net utility in the first period as a function of posterior belief x_1, anticipating optimal behavior in the second period

$$u_1(x_1) = \mathcal{V}(x_1) - k(x_1 - \mu)^2.$$

Therefore the manager's maximization problem in the two-period framework can be formalized using net utilities as follows

$$\begin{aligned} &\max_{p_1 \in \Delta[0,1]} \int_{[0,1]} u_1(x) dp_1(x), \\ \text{s.t.} \qquad &\mathcal{V}(x) = \max_{p_2 \in \Delta[0,1]} \int_{[0,1]} u_2(x, y) dp_2(y). \end{aligned} \tag{3.2}$$

3.3 Solution for the Two-Period Model

3.3.1 Solution for the Second Period

We solve the two-period model using backward induction, first solving the second-period problem (3.1). This problem is a variant of the binary static Matejka-McKay model with outside option with a different cost function (Matějka and McKay (2015)). Therefore, we claim that support of optimal distribution p_2 has at most two points.

One substantial difference in our setup is the possibility of full learning. Because the derivative of the cost function is bounded on the boundaries, full learning, in general, may be optimal. For simplicity, we only focus on the interior posterior distribution. We show that it is possible to eliminate full learning behavior by restricting the set of parameters.

Lemma 3.1. *If the prior belief of the manager is not too extreme, $\mu \in [\frac{1}{2k}, 1 - \frac{1}{2k}]$ for $k > 1$ then a full learning posterior belief $x_2 \in \{0, 1\}$ cannot be a part of the solution to problem (3.1).*

We consider the value of marginal cost k such that the interval for prior belief $[\frac{1}{2k}, 1 - \frac{1}{2k}]$ is substantively large.

We solve the problem (3.1) using the concavification technique. We use standard tangency conditions for the net utility at posterior realizations. Net utility depends on the prior belief and therefore optimal posteriors may also depend on the prior, which can complicate analysis. We show that, with quadratic cost this is not the case and, moreover, that this cost function specification allows us to find optimal posteriors and learning regions in closed form. The next proposition identifies optimal posteriors for all possible values of outside option x_1.[9]

Proposition 3.1. *If the value of the outside option (posterior realization from the first period) is too low, $x_1 \leq \mu - \frac{1}{4k}$ or too high, $x_1 \geq \mu + \frac{1}{4k}$ the manager does not learn (in the second period) and chooses the second candidate in the former case and the first candidate in the latter case.*

For intermediate values $x_1 \in (\mu - \frac{1}{4k}, \mu + \frac{1}{4k})$ the manager learns (in the second period). The optimal posterior distribution p_2 has two posterior beliefs, $x_2^L = x_1 - \frac{1}{4k}$ and $x_2^H = x_1 + \frac{1}{4k}$.

The solution has interval structure that is standard in the rational inattention literature. When the value of the outside option is too extreme with respect to the expected quality of a candidate μ, learning is too expensive, the manager does not consume any informative signals and chooses either the outside option or a candidate. In other cases, the manager learns and chooses an option depending on the realization of a signal.

Putting all variables together, we obtain a tractable expression for value function $\mathcal{V}(x_1)$, which is needed for the net utility in the first period. From the simple algebra we calculate that the value function in the learning region equals

$$\mathcal{V}(x_1) = kx_1^2 + ax_1 + b,$$

[9] This result is not entirely new. A variant of it has appeared in prior literature: see e.g. Jain and Whitmeyer (2021) and Wei (2021).

where $a = \frac{1}{2} - 2k\mu, b = k(\mu + \frac{1}{4k})^2$. Obviously $\mathcal{V}(x_1) = \mu$ if $x_1 \leq \mu - \frac{1}{4k}$ and $\mathcal{V}(x_1) = x_1$ if $x_1 \geq \mu + \frac{1}{4k}$. The function $\mathcal{V}(x_1)$ is clearly continuous on $[0, 1]$, and simple algebra shows that it is also continuously differentiable.

3.3.2 Solution for the First Period

Before analyzing the problem formally, we describe the trade-offs for the manager. Consider a distribution p_1 with two posterior realizations x_1^L, x_1^H. We interpret an interview, in which the manager asks several questions, which can be easy or hard. The expected value of a candidate does not change if he answers an easy question correctly or does not answer a hard one. If a candidate fails to answer an easy question, his expected quality decreases; if she succeeds on a hard one, his expected quality increases.[10]

Increasing x_1^H (decreasing x_1^L) is more costly for the manager because it increases the informativeness of an interview by adding one hard (easy) question. In the static problem, increasing x_1^H and decreasing x_1^L increases the manager's payoff simply by Blackwell's theorem. However, dynamics generates an additional effect on x_1^L. In the second period, the manager prefers to have a candidate with higher expected value x_1^L. This effect additionally discourages the manager from including an easy question in the interview. Thus, it is generally unclear which values of x_1^L, x_1^H are optimal for the manager.

Formally, in the solution the manager equates the marginal gains of x_1^L and x_1^H, that is $u_1'(x_1^L) = u_1'(x_1^H)$. Because function $\mathcal{V}(x_1)$ is (strictly) convex on the interval $(\mu - \frac{1}{4k}, \mu + \frac{1}{4k})$, it is generally also unclear on which interval optimal x_1^L and x_1^H belong.

However, having identified the functional form for $\mathcal{V}(x_1)$ analysis of problem (3.2) is straightforward. Because function $\mathcal{V}(x_1)$ is constant and linear on intervals $[0, \mu - \frac{1}{4k}]$ and $[\mu + \frac{1}{4k}, 1]$ correspondingly, net utility $u_1(x_1)$ is a concave quadratic polynomial on these intervals. Surprisingly, on the interval $[\mu - \frac{1}{4k}, \mu + \frac{1}{4k}]$ net utility is linear because quadratic terms cancel out.

Because the function $\mathcal{V}(x_1)$ is continuously differentiable, the net utility $u_1(x_1)$ is also continuously differentiable. Therefore, the net utility is weakly concave on the $[0, 1]$ and

[10] Strictly speaking, this interpretation of an interview is not formally correct. Because the manager is Bayesian, adding one more question provides information in success and failure cases. Therefore, in this case both x_1^H and x_1^L should change: x_1^H increases and x_1^L decreases. In our interpretation, we use the following approximation. Our interpretation of a hard (easy) question is that the information in the case of a failure (success) is so little that it could be neglected. In more formally correct interpretation, when, for example, only x_1^H increases and x_1^L stays fixed, the manager slightly restructures the interview by, for example, making one of the questions a bit harder. We apply the former interpretation primarily for brevity.

its concavification coincides with the function itself. The prior belief μ lies on the linear region $[\mu - \frac{1}{4k}, \mu + \frac{1}{4k}]$ therefore, the solution to the problem is not unique. The necessary and sufficient condition for optimal feasible posterior beliefs is to lie inside interval $[\mu - \frac{1}{4k}, \mu + \frac{1}{4k}]$. For convenience, we formalize our result with posterior distribution, which has at most two posterior beliefs.

Theorem 3.1. *Suppose that the posterior distribution p_1 has at most two posterior realizations. In the solution to problem (3.2) in the first period, it is optimal either not to learn anything or to have posterior distribution p_1 with two posterior beliefs, $x_1^L \in [\mu - \frac{1}{4k}, \mu)$ and $x_1^H \in (\mu, \mu + \frac{1}{4k}]$.*

Optimal posterior distribution p_1 is not unique. Moreover, in the solution to problem (3.2), any non-degenerate posterior belief $x_1 \in (\mu - \frac{1}{4k}, \mu + \frac{1}{4k})$ leads to learning in the second period. Additionally, the manager chooses both candidates with positive probabilities given any optimal first-period posterior belief x_1. The recommendation lemma, which suggests a one-to-one mapping between posterior belief and action[11], does not hold in our setup. The main formal argument for the recommendation lemma is the linearity of function $\mathcal{V}(x_1)$ in x_1. In our case, function $\mathcal{V}(x_1)$ is clearly non-linear in x_1.

To analyze how the order of candidates leads to discrimination, we consider how unconditional choice probabilities change with changing learning strategy. We fix symmetric first period posteriors $x_L^1 = \mu - \Delta, x_H^1 = \mu + \Delta$ and allow the manager to obtain a more precise high posterior belief $x_H^1 = \mu + \Delta + \varepsilon$. The manager chooses the first candidate if she receives a low signal in the second period. In the second period, given the value of the outside option equals x_1 the manager chooses the first candidate with probability $\frac{1}{2} + 2k(x_1 - \mu)$. Therefore, the unconditional probability of choosing the first candidate equals

$$P(\text{choose } 1) = \left(\frac{1}{2} - 2k\Delta\right)\frac{\Delta + \varepsilon}{2\Delta + \varepsilon} + \left(\frac{1}{2} + 2k(\Delta + \varepsilon)\right)\frac{\Delta}{2\Delta + \varepsilon} = \frac{1}{2}.$$

Intuitively, increasing x_1^H corresponds to asking one more hard question. This generates two effects. First, the first interview becomes harder and the first candidate's probability of success decreases by $2k\Delta\frac{\Delta+\varepsilon}{2\Delta+\varepsilon}$. Second, if the first candidate succeeds, the manager constructs a harder interview in the second stage. In this case, the second candidate fails more often and the probability of choosing the first candidate increases

[11]For the formal statement with entropy cost, see Lemma 1 in Matějka and McKay (2015)

by $2k(\Delta+\varepsilon)\frac{\Delta}{2\Delta+\varepsilon}$. Because $2k\Delta\frac{\Delta+\varepsilon}{2\Delta+\varepsilon} = 2k(\Delta+\varepsilon)\frac{\Delta}{2\Delta+\varepsilon}$ holds, the average probability of choosing the first candidate does not change.

Clearly, in the solution, the manager can learn differently in different periods. For example, Theorem 3.1 suggests that the strategy in which the manager does not learn in one period and chooses a posterior distribution with $x_L = \mu - \frac{1}{4k}, x_H = \mu + \frac{1}{4k}$ in another is optimal. Obviously, on aggregate, the manager obtains the same amount of information in the two strategies. This observation holds for any optimal learning strategy. We show that, in an optimum strategy, the manager always obtains a fixed amount of information $\bar{c}$. She may incur any amount of information below $\bar{c}$ in the first period and the rest in the second. Therefore, the manager perfectly substitutes information between periods. The following corollary describes the properties of the solution to problem (3.2).

Corollary 3.1. *In the solution to problem (3.2) the manager chooses candidates with equal unconditional probabilities. In all optimal strategies, the manager obtains the fixed amount of information $\bar{c}$: she may obtain any non-negative amount in the first period and the rest in the second.*

The manager may learn more in the first period, or in the second; therefore the primacy or recency effect may present. The exact conditions for a serial-position effect are not identifiable. Everything goes: the manager is free to obtain any amount of information below $\bar{c}$ in the first period.

The manager does not discriminate against a candidate on average. The amount of information the manager obtains does not correlate with the unconditional probability of choice. Even if the manager obtains a higher amount of information during her interview with candidate i, she chooses him with probability $\frac{1}{2}$. However, the manager may discriminate against a candidate conditionally. For example, if she learns only in the first period, she clearly chooses the first candidate conditional on him being a high type more often, because $x_1^H > \mu$.

In the next section, we show how including an additional alternative changes our results: the manager discriminates against the last candidate and the primacy effect is present.

3.4 An Additional Alternative

In this section, we show how including an additional alternative creates discrimination in the choice. When three alternatives are available to the manager, she acquires information differently, and perfect substitutability of the information also does not hold. For an exposition of our results, it is sufficient to extend the setup by adding an additional candidate with intermediate expected quality $x_1 \approx \mu$ and to not allow the manager to learn about this candidate. We keep the structure of the problem the same and investigate a slightly richer model. We analyze a model in which the manager inspects three identical candidates sequentially. We define the manager's problem formally, and similarly to the two-period problem, we introduce the manager's problem sequentially.

We fix the first and second-period learning strategies and denote them as p_1, p_2. After the first two periods, the manager obtains posterior beliefs x_1, x_2 about the qualities of the candidates. These realizations serve as the outside options of the manager. We define the net utility function in the third period given x_1, x_2 and the manager's posterior belief about the third candidate x_3 as

$$u_3(x_1, x_2, x_3) = \max\{x_1, x_2, x_3\} - k(x_3 - \mu)^2.$$

Therefore the third period problem is

$$\max_{p_3 \in \Delta[0,1]} \int_{[0,1]} u_3(x_1, x_2, x) dp_3(x). \tag{3.3}$$

for feasible distribution p_3. We denote the maximum attained value in the problem as $\mathcal{V}_3(x_1, x_2)$. For the second period, we similarly define net utility as a function of posterior belief x_1, anticipating optimal behavior in the third period

$$u_2(x_1, x_2) = \mathcal{V}_3(x_1, x_2) - k(x_2 - \mu)^2$$

Therefore the second period problem of the manager is

$$\begin{aligned} &\max_{p_2 \in \Delta[0,1]} \int_{[0,1]} u_2(x_1, x) dp_2(x), \\ &\text{s.t.} \qquad \mathcal{V}_3(x_1, x) = \max_{p_3 \in \Delta[0,1]} \int_{[0,1]} u_3(x_1, x, y) dp_3(y). \end{aligned} \tag{3.4}$$

Finally, if the maximum attained value in the problem above equals $\mathcal{V}_2(x_1)$ then the first period net utility equals

$$u_1(x_1) = \mathcal{V}_2(x_1) - k(x_1 - \mu)^2$$

and the manager's problem in the three-period setup can be formalized as

$$\begin{aligned} &\max_{p_1 \in \Delta[0,1]} \int_{[0,1]} u_1(x) dp_1(x), \\ \text{s.t.} \qquad &\mathcal{V}_2(x) = \max_{p_2 \in \Delta[0,1]} \int_{[0,1]} u_2(x,y) dp_2(y). \end{aligned} \tag{3.5}$$

3.4.1 Solution to the Model With Three Alternatives

Denoting $x_{1,2} = \max\{x_1, x_2\}$ it is clear that the problem in the third period (3.3) is equivalent to the second-period problem (3.1) with posterior belief $x_{1,2}$ after the first period. Therefore the solution to problem (3.3) follows from Proposition 3.1.

Problem (3.4) differs from problem (3.2) due to the presence of the outside option from the first period x_1. We discuss the properties of the solution to the problem (3.4) in the appendix. Formally, the value function $\mathcal{V}_2(x)$ has an interval structure as in problem (3.2), but is non-linear in the intermediate region. Because of this non-linearity, the solution to the three-alternative problem may seem complicated. However, we show that it has a simple structure.

Theorem 3.2. *In the model with three alternatives, at stage i, it is optimal either to have a distribution p_i with two posterior beliefs x_i^L, x_i^H or not to consume any informative signals. The manager does not consume informative signals at stage i given high posterior realization in previous stages and chooses p_i otherwise.*

If the manager receives high posterior realization in stage i, she chooses candidate i. If the manager does not receive any high posterior realizations, she chooses the first candidate. Additionally, inequalities

$$x_3^L < x_2^L < x_1^L < x_3^H < x_2^H < x_1^H$$

hold.

The manager continues an active search only if she receives a low signal and chooses an alternative otherwise. This pattern resembles the behavior from Simon (1955)'s classical model of satisficing . However, our framework creates a generalization of the satisficing behavior. Unlike Simon (1955), in our model, the choice threshold decreases over time. Interestingly, this result is driven only by the presence of endogenous information, but not, for example, by discounting. Decreasing threshold behavior is optimal in a variant of the "secretary problem" (Gilbert and Mosteller (1966)). In the "secretary problem", the decision maker samples i.i.d. items and decides when to stop searching and to choose an option. In the "secretary problem", the search is costless, and the discarded option cannot be chosen in later stages; therefore, the driver of our result is different.

In our analysis in this section, we compare the unconditional probabilities of the choice of the second or third candidates. There are two main reasons for this comparison. First, in our analysis above, we investigate order effects in the framework with two candidates. Therefore, the natural interpretation of the first candidate in the current framework is an endogenous outside option. Second, we find conditions for optimal information in the first period and the choice probabilities of choosing the first candidate to be generally analytically intractable.

To investigate the impact of the additional candidate on the second and the third candidates, we analyze problem (3.4) with an outside option x_1^L. From the proof of Theorem 3.2 the inequality for the low posterior belief $x_1^L \in (\mu - \frac{1}{4k}, \mu)$ holds. Therefore, useful to study how the solution to problem (3.4) changes for different values of x_1^L on the interval $[\mu - \frac{1}{4k}, \mu]$. Our analysis of the problem is based on the proof of Theorem 3.2, which appears in the appendix.

If the value of an outside option is too low $x_1 = \mu - \frac{1}{4k}$, it does not make a difference to the manager's choice, and the solution to problem (3.4) is the same as in the model with two candidates. For greater values of x_1^L, the manager uses this posterior as the best option in the case of two low realizations in the future. In the second stage, outside option x_1^L incentivizes the manager to choose p_2 such that $x_2^H > \mu + \frac{1}{4k}$.

To illustrate the different predictions from Theorem 3.1 and Theorem 3.2, we fix the optimal learning strategy in the current setup, in which p_2 has posterior beliefs x_2^L, x_2^H. The strategy that leads to the same posteriors is available in the case with two candidates. With two candidates, to have the same posteriors, the manager should choose p_1 with posterior beliefs x_1^L, x_2^H and should behave optimally in the second period. However, this strategy is not optimal. Because $x_1^L > x_2^L$, the first interview provides too little information to the manager. Intuitively, she asks too few easy questions during the interview.

Intuitively, in the second period, the manager optimally balances between the informativeness and the difficulty of the interview. She chooses x_2^H to be sufficiently high that it is sufficient not to consume informative signals in the third stage. She chooses x_2^L to be sufficiently low to decrease the overall difficulty of the interview and increase its informativeness. Such a careful design generates discrimination against the third candidate, and the manager unconditionally chooses the second candidate more often than the third.

Because inequalities $x_2^H > x_3^H, x_2^L > x_3^L$ hold, intuitively, in the third stage, the manager constructs a simpler interview: she asks more easy questions and fewer hard questions. Moreover, because $x_2^H > x_3^H$ the manager engages in "cherry-picking" behavior. Bartoš et al. (2016) find similar behavior in their field experiment in an environment with information acquisition.

Additionally, we relate the discrimination in choice with the serial-position effect. We show that the manager, on aggregate, obtains a higher amount of information during the second interview, and, therefore, the primacy effect is present. The manager obtains a higher amount of information in the second period mainly because she does not consume informative signals in the third stage after receiving posterior x_2^H. Indeed, we show that $C(p_2) < C(p_3)$ holds. The manager wants to postpone the information acquisition to the third stage. She designs the harder interview in the second period, which is less costly. The next corollary summarizes the discussion above.

Corollary 3.2. *In the model with three alternatives, the manager, on aggregate, obtains a higher amount of information in the second period than in the third. Also, the manager unconditionally chooses the second candidate more often than the third one.*

3.5 Extensions

This section considers simple generalizations of our frameworks with two and three candidates. With three candidates for simplicity, we consider a variant in which the first candidate has an exogenous quality $x_1 \approx \mu$ and we analyze the manager's behavior in the second and the third periods. In this section, we study a setup in which candidates differ in terms of expected quality. We also introduce discounting and consider different cost function specifications. In all extensions, we keep the assumption on the interior learning, $\mu \in [\frac{1}{2k}, 1 - \frac{1}{2k}], k > 1$.

3.5.1 Different Candidates

We start with a setup in which candidates have different expected qualities. We consider problem (3.2) and assume that the first candidate has expected quality μ_1 and the second has μ_2. Our main technical argument in Theorem 3.1 deals with the behavior of the marginal utility therefore, unsurprisingly, a similar result holds in this setting.

Lemma 3.2. *If $|\mu_1 - \mu_2| \leq \frac{1}{4k}$ then it is optimal in the first period either not to learn anything or to have posterior distribution p_1 with two posterior beliefs, $x_1^L \in [\mu_2 - \frac{1}{4k}, \mu_1)$ and $x_1^H \in (\mu_1, \mu_2 + \frac{1}{4k}]$.*

As in Theorem 3.1, for simplicity, we formulate our result with posterior distribution, which has at most two posterior beliefs. One additional requirement is needed for Lemma 3.2 to hold: candidates should not differ much in their expected qualities. Technically, in this case, the prior belief in the first period belongs to the learning region.

In the current setup, similar optimal learning behavior leads to qualitatively similar predictions. The unconditional probability of choice and the total amount of information obtained are constant across all optimal learning strategies. We formulate the properties of the solution in the corollary below.

Corollary 3.3. *Fix $\mu_1 \in (\mu_2, \mu_2 + \frac{1}{4k})$. The manager chooses a first candidate with unconditional probability $\frac{1}{2} + 2k(\mu_1 - \mu_2)$ and the second candidate with unconditional probability $\frac{1}{2} + 2k(\mu_2 - \mu_1)$. The manager obtains the fixed amount of information $\bar{c} - k(\mu_1 - \mu_2)^2$, and she may obtain any non-negative amount in the first period and the rest in the second.*

The manager unconditionally chooses a better candidate more often. As in the solution to problem (3.1), the manager uses the information as a perfect substitute in different periods. The exact conditions for the primacy or recency effect are not identifiable, and the amount of information obtained is unrelated to the probability of choice. Additionally, the manager obtains a lesser amount of information on aggregate than in problem (3.1). This is intuitively true because the information is most valuable in the least certain environment, which is the formulation with identical candidates.

We finish the analysis of the model with two different candidates with the optimal order question. Because the candidates differ, the manager may benefit from the particular inspection order. For example, Weitzman (1979) suggests that the manager should start her inspection with the better candidate. We show that, in our case, the manager is indifferent between different order of inspection.

Corollary 3.4. *If $|\mu_1 - \mu_2| \leq \frac{1}{4k}$ then the manager obtains the same value in the problem no matter which candidate she inspects first.*

If values μ_2 and μ_3 are sufficiently close to each other, including an additional alternative with intermediate exogenous quality x_1 leads to a similar optimal learning strategy as in problem (3.5). The optimal manager's strategy preserves the same features: the manager chooses distribution p_2 with posterior beliefs x_2^L, x_2^H such that she does not consume any informative signals after realization x_2^H. In general, it may happen that although the second candidate has a lower expected quality $\mu_2 < \mu_3$, the manager chooses him more often. For example, it is true if $k = 2$ and $x_1 = 0.5, \mu_3 = 0.5, \mu_2 = 0.49$. Intuitively, because optimal x_2^L, x_2^H does not depend on μ_2, the lower expected quality μ_2 does not change the structure of the interview. If μ_2 slightly decreases, the second candidate succeeds less often. However, if μ_2 is close to μ_3, discrimination against the third candidate remains. We show that such behavior is never optimal if the manager also chooses the order of candidates.

If the manager optimally chooses the order of candidates, we show that she inspects the better candidate first and chooses him more often. Recall that the manager chooses more difficult interview questions in the second stage. Intuitively, it is efficient for her to ask harder questions to a better candidate. We formally show the result when μ_2 and μ_3 are marginally different and summarize the properties of the solution to the model with three alternatives in the corollary.[12]

Corollary 3.5. *$\delta_\mu > 0$ exists such that, if $|\mu_2 - \mu_3| < \delta_\mu$ then in the second stage the manager chooses p_2 with two posterior beliefs x_2^L, x_2^H. She consumes an informative signal in the third stage only given x_2^L realization.*

$\varepsilon_\mu > 0$ exists such that, if $\mu_i \in (\mu_j, \mu_j + \varepsilon_\mu)$, the manager inspects candidate i in the second stage and chooses him unconditionally more often.

3.5.2 Discounting

To fully capture the sequential effect of the problem in our prior analysis, we assumed that the manager valued information to the same degree in different periods. This is

[12]We do not study the serial-position effect in this setting because the cost of learning in stage i depends on μ_i. This makes direct comparison of the amount of information in different stages ambiguous.

clearly an essential simplifying assumption, but we relax it in this section. The most natural way to do is by introducing discounting, because it is widely documented that people value present events more than future ones.

In this section, we assume that the manager discounts only the information but not the expected value of a candidate. There are two main reasons for this. Firstly, our goal is to identify the interaction only between discounting of information and the sequential nature of the problem. Secondly, the expected value of a candidate can be seen as his long-run average productivity and, therefore, should not be discounted.

We start with a problem with two candidates (3.1). That is, we assume that the marginal cost of information differs in different periods, $k_1 = tk_2, t > 1$. Basic intuition suggests that, with discounting, the information in the second period becomes cheaper, and the intertemporal role of the information should change. We show that the basic intuition holds. Moreover, the manager learns only in the second period.

Corollary 3.6. *In the solution to problem (3.2) with different marginal costs of information $k_1 > k_2$ the manager does not learn in the first period. In the second period she chooses p_2 such that $x_2^L = \mu - \frac{1}{4k}$, $x_2^H = \mu + \frac{1}{4k}$.*

Discounting does not essentially change the results of Theorem 3.1. The manager does not discriminate between candidates and chooses them uniformly, on average. Moreover, the manager obtains the same amount of information as in the setup without discounting. Because the information is more expensive in the first period, the manager simply obtains it only in the second period, and the recency effect occurs. The amount of information the manager obtains does not relate to discrimination, on average. Because $x_2^H > \mu$, she chooses the second candidate more often, given that he is a high-quality candidate.

If t is low enough, discounting does not qualitatively change the optimal behavior of the manager with three alternatives: she chooses p_2 with posterior beliefs x_2^L, x_2^H such that she does not consume any informative signals after realization x_2^H.

However, discounting generates an interesting effect on the relationship between discrimination and information obtained. For simplicity, we consider the case $\mu = x_1$. In this case, the difference in information obtained between periods is perfectly correlated with the discrimination against a candidate. Suppose that k_2 is sufficiently close to k_3. In that case, the results are the same as without discounting: the manager obtains more information in the second period and chooses the second candidate more often. However, if k_2 is high enough, the information becomes much cheaper in the third period, and the manager reflects this in her solution. She obtains more information about the third candidate and chooses the second candidate the least often.

If k_2 increases, the manager obtains less information in the second period in the Blackwell sense. Intuitively, the second interview becomes too difficult, and the second candidate fails so often that the manager chooses him with the least unconditional probability. We summarize the properties of the solution to the model with three alternatives in the corollary below.

Corollary 3.7. $\delta_k > 1$ *exists such that, if* $t \in (1, \delta_k)$ *then in the second stage the manager chooses* p_2 *with two posterior beliefs* x_2^L, x_2^H*. She consumes an informative signal in the third stage only given* x_2^L *realization.*

Fix $x_1 = \mu$*.* $\bar{t}, \underline{t}$ *with* $\bar{t} > \underline{t} > 1$ *exist such that, if* $k_2 \in (k_3, \underline{t}k_3)$ *the manager unconditionally chooses the second candidate more often and on average obtains more information in the second period. If* $k_2 \in (\underline{t}k_3, \bar{t}k_3)$ *the manager unconditionally chooses the third candidate more often and, on average, obtains more information in the third period.*

3.5.3 Different Cost Function

Technically, our results depend heavily on the quadratic cost function, therefore it is important to consider different cost specifications. In this section, we keep the power functional form assumption and consider costs to be equal

$$\tilde{C}(p) = k \int_{[0,1]} (x - \mu)^4 dp(x)$$

for the chosen distribution p. The function $\tilde{C}(p)$ is also the posterior-separable cost function.[13] Additionally, the net cost of posterior belief x is the convex transformation of the quadratic cost from the main model, therefore, any non-degenerate posterior distribution p is cheaper with cost $\tilde{C}(p)$.

We found the model with cost function $\tilde{C}(p)$ analytically intractable, so we primarily discuss whether our main results extend using numerical computations. Our analyses suggest that the no-discrimination result holds with two identical candidates, but the perfect substitutability of information does not. For example, for $k = 4$ and $\mu \in (\frac{2}{5}, \frac{3}{5})$, the manager obtains more information in the third period, and the recency effect is

[13] An additional special feature of the power functional form cost is its symmetry. Because such a form is based on the distance between the posterior and the prior, acquiring positive and negative information is equally costly for the decision-maker.

present. As with quadratic cost, the amount of acquired information does not lead to discrimination in the choice. Interestingly, the difference in information obtained in two periods is constant over some interval of μ, which suggests some existence of the underlying structure for information obtained in the optimum.

In the setup with the three alternatives, the current cost specification generates similar qualitative results as in the case with quadratic costs. In the figure below, we present an example of an optimal learning strategy in the second period for $k = 4, \mu = x_1 = \frac{1}{2}$. In the second period, the manager chooses distribution p_2 with two realizations such that she consumes an informative signal only after receiving x_2^L. Moreover, discrimination against the third candidate relates to the primacy effect. Unconditionally, the manager chooses the second candidate more often as $P(\text{choose } 2) \approx 0.43, P(\text{choose } 3) \approx 0.28$ and on average obtains more information in the second period, as $\tilde{C}(p_2) \approx 0.029, \tilde{C}_3(p_2) \approx 0.022$.

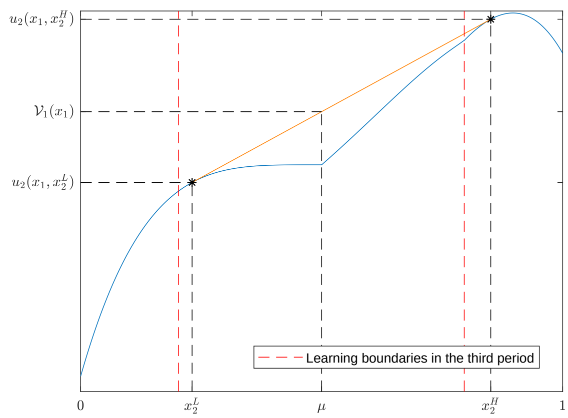

Figure 3.1: The solution to the setup with three alternatives for $k = 4, \mu = x_1 = \frac{1}{2}$ and $\tilde{C}(p)$ cost.

A model with the current cost specification generates qualitatively similar results that suggest that our findings hold in a more general setting. We leave the investigation of this for future research.

3.6 Conclusion

We consider a stylized model of interview design in which the manager learns about candidates sequentially and flexibly. We study the relationships between the order of inspection and discrimination in choice and connect discrimination to the serial-position effect. Our findings vary depending on the number of available candidates. In a case with two candidates, there is no discrimination, and the amount of information the manager obtains is not connected to the choice. However, with three alternatives the manager discriminates against the last candidate and the primacy effect is present. She obtains the least amount of information on average during the last interview. Our results are robust to the simple generalizations of our model and also vary depending depending on the number of candidates.

Generalizing the cost function in our main model is a natural direction for future research. We chose the quadratic cost mostly for tractability, and our numerical results with a different cost function suggest that our results are robust for the general class of cost functions. Additionally, it would be interesting to identify which property of the cost function leads to the primacy and recency effects.

Another important extension of our model worth developing is expanding the number of candidates. It is interesting to identify whether the framework with two candidates is a special case and whether the structure of an interview resembles our findings in the case with three candidates. For example, does the difficulty of an interview decrease in the order of candidates, and does the manager engage in "cherry-picking" behavior? The bottleneck of this research direction is finding an analytically tractable setup that allows the study of discrimination and its relation to the serial-position effect.

3.A Proofs

3.A.1 Proof of Proposition 3.1

To solve the problem, we use concavification analysis, as in, e.g., Caplin and Dean (2013). We consider net utility $u_2(x_1, x_2)$ as a function of x_2 and find its concave closure. Because x_2 is a posterior belief, $x_2 \in [0, 1]$ holds. We first consider the function on the whole real line and find its concave closure. The function is concave on intervals $(-\infty, x_1)$ and (x_1, ∞), but not jointly because the $\lim_{x_2 \to x_1 - 0} \frac{\partial u_2(x_1,x_2)}{\partial x_2} < \lim_{x_2 \to x_1 + 0} \frac{\partial u_2(x_1,x_2)}{\partial x_2}$.

To obtain the concave closure of function $u_2(x_1, x_2)$ we need to connect the part to the left of x_1 and to the right of x_1 with a straight line. Because the function is differentiable, the tangency condition holds. We denote these tangency points as z' and z''. The concave closure equals $u_2(x_1, x_2)$ on intervals $(-\infty, z')$ and (z'', ∞), and is a straight line, which connects points $(z', u_2(x_1, z'))$ and $(z'', u_2(x_1, z''))$ on interval $[z', z'']$.

This approach leads to the concave closure by the standard arguments: the function is concave by construction; it lies weakly above function $u_2(x_1, x_2)$ because a concave function lies below its tangent line; and the resulting function is clearly the minimal concave function.

Clearly, if $z', z'' \in [0, 1]$, then these points are optimal two posterior beliefs. We denote them as x_2^L, x_2^H, where $x_2^L < x_1$ and $x_2^H > x_1$. Two necessary and sufficient conditions for these points: the concave closure has the same slope at points x_2^L, x_2^H; this slope equals the ratio between $u_2(x_1, x_2^H) - u_2(x_1, x_2^L)$ and $x_2^H - x_2^L$.

The first condition results in:

$$\frac{\partial u_2(x_1, x_2^H)}{\partial x_2} = \frac{\partial u_2(x_1, x_2^L)}{\partial x_2} \quad \Leftrightarrow \quad 1 - 2k(x_2^H - \mu) = -2k(x_2^L - \mu) \quad \Leftrightarrow \quad x_2^H - x_2^L = \frac{1}{2k}.$$

The second condition results in:

$$\frac{\partial u_2(x_1, x_2^L)}{\partial x_2} = \frac{u(x_1, x_2^H) - u(x_1, x_2^L)}{x_2^H - x_2^L}$$

Because

$$u(x_1, x_2^H) - u(x_1, x_2^L) = \mu - x_1 + \frac{1}{4k}$$

the second condition becomes

$$-2k(x_2^L - \mu) = 2k(\mu - x_1 + \frac{1}{4k}) \quad \Leftrightarrow \quad x_2^L = x_1 - \frac{1}{4k}, x_2^H = x_1 + \frac{1}{4k}.$$

If

$$\mu < x_2^L \quad \Leftrightarrow \quad x_1 > \mu + \frac{1}{4k}$$

then point $(\mu, u_2(x_1, \mu))$ lies in the concave region, the manager does not learn, and choose the first candidate. If

$$\mu > x_2^H \quad \Leftrightarrow \quad x_1 < \mu - \frac{1}{4k}$$

then point $(\mu, u_2(x_1, \mu))$ lies in the concave region, the manager does not learn, and choose the second candidate. If

$$\mu \in [x_2^L, x_2^H] \quad \Leftrightarrow \quad x_1 \in [\mu - \frac{1}{4k}, \mu + \frac{1}{4k}]$$

then then manager learns and chooses posterior distribution with posterior beliefs x_2^L and x_2^H.

3.A.2 Proof of Lemma 3.1

We analyze problem (3.1) for such values x_1 and k that optimal interior posteriors do not lie inside the $[0,1]$ interval.

Clearly, if x_1 is close to 0 or 1, optimal interior posteriors $x_1 - \frac{1}{4k}, x_1 + \frac{1}{4k}$ do not lie in the $[0,1]$. The necessary condition for two interior posteriors to lie in $[0,1]$ is to hold

$$x_1 - \frac{1}{4k} > 0, \quad x_1 + \frac{1}{4k} < 1,$$

where $x_1 = \frac{1}{4k} + \varepsilon$ for small enough ε. The inequalities above result in inequality $k > \frac{1}{2}$. Therefore if $k \leq \frac{1}{2}$, then for any posterior value x_1, interior learning in the second period is not possible. We eliminate this possibility in our analysis and will always consider a case with $k > \frac{1}{2}$.

If $k > \frac{1}{2}$ then at least one interior posterior lies inside $[0,1]$. This holds because in this case two inequalities $x_1 < \frac{1}{4k}$ and $x_1 > 1 - \frac{1}{4k}$ cannot hold jointly.

We start with the case in which the low posterior lies outside of the $[0,1]$ interval: $x_1 - \frac{1}{4k} < 0$ and $x_1 + \frac{1}{4k} < 1$.

We claim that the concave closure is a straight line that connects points $(0, u_2(x_1, 0))$ and $(z'', u_2(x_1, z''))$, which is tangent to $u_2(x_1, x_2)$ at point $(z'', u_2(x_1, z''))$, and equals $u_2(x_1, x_2)$ on interval $[z'', 1]$. Intuitively, the manager wants to learn more about state $\theta_2 = 0$ but faces the boundary constraint.

We denote a straight line from above as $y = \alpha x_2 + \beta$. From the analysis above $x_2^L = 0$ and three conditions for x_2^H are

$$\begin{cases} \alpha = 1 - 2k(x_2^H - \mu), \\ \beta = x_1 - k\mu^2, \\ x_2^H - k(x_2^H - \mu)^2 = \alpha x_2^H + \beta \end{cases}$$

Simple algebra shows that the high posterior is $x_2^H = \sqrt{\frac{x_1}{k}}$.

The constructed function is the concave closure by standard arguments. It is concave by construction. Because inequality $\sqrt{\frac{x_1}{k}} < \mu + \frac{1}{4k}$ holds, inequality $\frac{\partial u_2(x_1,0)}{\partial x_2} < \alpha$ also holds. Therefore the constructed function is weakly greater than $u_1(x_1, x_2)$ on all $x_2 \in [0, 1]$. It is also clearly the minimal concave function.

We continue the analysis with the case in which only the high interior posterior lies outside of the $[0, 1]$ interval: $x_1 - \frac{1}{4k} > 0$ and $x_1 + \frac{1}{4k} > 1$.

The analysis is very similar to the previous case. From the same arguments as above the concave closure on the $[0, z']$ interval equals function $u_2(x_1, x_2)$, and on the $(z', 1]$ the closure is a straight line that connects points $(z', u_2(x_1, z'))$ and $(1, u_2(x_1, 1))$. Intuitively, the manager wants to learn more about state $\theta_2 = 1$ but faces the boundary restriction.

We denote a straight line from concave closure as $y = \alpha x_2 + \beta$. From the analysis above, three conditions for x_2^L are

$$\begin{cases} \alpha = -2k(x_2^L - \mu), \\ \alpha + \beta = 1 - k(1 - \mu)^2, \\ x_1 - k(x_2^L - \mu)^2 = \alpha x_2^L + \beta \end{cases}$$

Simple algebra shows that the low posterior is $x_2^L = 1 - \sqrt{\frac{1-x_1}{k}}$.

Therefore, full learning posterior belief $x_2 \in \{0, 1\}$ is a part of the solution only in two cases: if the outside option and prior belief are low enough $x_1 < \frac{1}{4k}, \mu < \sqrt{\frac{x_1}{k}}$ or if the outside option and prior belief are high enough $x_1 > 1 - \frac{1}{4k}, \mu > 1 - \sqrt{\frac{1-x_1}{k}}$. Thus, if μ belongs either $[0, \frac{1}{2k}]$ or $[1 - \frac{1}{2k}, 1]$ then the full learning posterior may be optimal. Therefore, if $\mu \in (\frac{1}{2k}, 1 - \frac{1}{2k})$ then a full learning posterior can not be a part of the solution to problem (3.1).

3.A.3 Proof of Corollary 3.1

The proof of the corollary mostly follows from simple algebra.

We denote posteriors in the first periods as x_1^L and x_1^H. The probability to choose the second candidate equals

$$P(\text{choose } 2) = \frac{x_1^H - \mu}{x_1^H - x_1^L} 2k(-x_1^L + \frac{1}{4k} + \mu) + \frac{\mu - x_1^L}{x_1^H - x_1^L} 2k(-x_1^H + \frac{1}{4k} + \mu).$$

From the simple algebra above equals $\frac{1}{2}$.

Given a posterior distribution p_2 with two posteriors realizations $x_2^1 < x_2^2$ the cost of learning equals

$$C(p_2) = k\Big(\frac{x_2^2 - \mu}{x_2^2 - x_2^1}(\mu - x_2^1)^2 + \frac{\mu - x_2^1}{x_2^2 - x_2^1}(\mu - x_2^2)^2\Big) = k(x_2^2 - \mu)(\mu - x_2^1).$$

Therefore given the posterior belief in the first period, x_1 cost of optimal learning in the second period equals

$$C(p_2(x_1)) = k(x_1 + \frac{1}{4k} - \mu)(\mu - x_1 + \frac{1}{4k}) = -k(x_1 - \mu)^2 + \frac{1}{16k}$$

Given the optimal posterior distribution p_1 with two posterior realizations $x_1^1 < x_1^2$ the aggregate cost of learning in the second period equals

$$C_2(p_1) = \frac{x_1^2 - \mu}{x_1^2 - x_1^1} C(p_2(x_1^1)) + \frac{\mu - x_1^1}{x_1^2 - x_1^1} C(p_2(x_1^2))$$

From the simple algebra

$$C_2(p_1) = -C(p_1) + \frac{1}{16k}.$$

holds. If the manager does not learn in the first period, then in the second, she chooses $x_2^L = \mu - \frac{1}{4k}, x_2^H = \mu + \frac{1}{4k}$ and obtains $\frac{1}{16k}$ amount of information in the second period. If the manager obtains the most information in the first period $\frac{1}{16k}$ and chooses p_1 such that $x_1^L = \mu - \frac{1}{4k}, x_1^H = \mu + \frac{1}{4k}$ then she does not learn in the second period. If the manager chooses p_1 optimally such that $C(p_1) \in (0, \frac{1}{16k})$ then she chooses p_2 optimally such that $C_2(p_1) = \frac{1}{16k} - C(p_1)$.

3.A.4 Proof of Lemma 3.2

The proof is based on the proof of Theorem 3.1. In the solution to problem (3.1), the optimal posteriors do not depend on the prior belief. Therefore the only condition that is needed for interior posteriors in the optimum is $\mu_2 \in (\frac{1}{2k}, 1 - \frac{1}{2k})$.

In problem (3.2) with different prior beliefs, function $u_1(x_1)$ is concave by construction and is linear in the intermediate region if $x_1 \in [\mu_2 - \frac{1}{4k}, \mu_2 + \frac{1}{4k}]$. Therefore the optimality condition is simply $\mu_1 \in [\mu_2 - \frac{1}{4k}, \mu_2 + \frac{1}{4k}]$. This condition is equivalent to $|\mu_1 - \mu_2| \leq \frac{1}{4k}$.

3.A.5 Proof of Corollary 3.3

The proof is identical to the proof of Corollary 3.1. It follows from the simple algebra replacing the value of the prior belief in different periods as μ_1 and μ_2.

3.A.6 Proof of Corollary 3.4

The proof follows from the simple algebra. We denote the value to problem (3.2) as $F(\mu_i, \mu_j)$ if in the first period the prior equals μ_i and in the second equals μ_j.

From the simple algebra

$$F(\mu_1, \mu_2) = \mu_1(\frac{1}{2} - 2k(-\mu_1 + \mu_2)) + k(\mu_2 + \frac{1}{4k})^2 - k\mu_1^2$$

holds. Taking the difference leads to

$$F(\mu_1, \mu_2) - F(\mu_2, \mu_1) = 0.$$

Thus the manager reaches the same maximum level of utility irrespectively of the candidate's order.

3.A.7 Proof of Corollary 3.6

The proof is based on the proof of Theorem 3.1. We need to analyze the derivative of the function $u_1(x)$. Simple algebra show that

$$u_1'(x) = \begin{cases} 2k_1(\mu - x) & \text{if } x < \mu - \frac{1}{4k_2} \\ \\ 2(k_2 - k_1)x + \frac{1}{2} - 2k_2\mu + 2k_1\mu & \text{if } x \in [\mu - \frac{1}{4k_2}, \mu + \frac{1}{4k_2}] \\ \\ 1 + 2k_1(\mu - x) & \text{if } x > \mu + \frac{1}{4k_2} \end{cases}$$

holds. Function $u_1'(x)$ decreases from $2k_1\mu$ to $\frac{k_1}{2k_2}$ on $[0, \mu - \frac{1}{4k_2})$; decreases from $\frac{k_1}{2k_2}$ to $1 - \frac{k_1}{2k_2}$ on $[\mu - \frac{1}{4k_2}, \mu + \frac{1}{4k_2}]$; decreases from $1 - \frac{k_1}{2k_2}$ to $1 + 2k_1(\mu - 1)$ on $(\mu + \frac{1}{4k_2}, 1]$. Therefore $u_1'(x)$ decreases on $[0, 1]$ and function $u_1(x)$ is concave. Therefore, the manager does not consume informative signals in the first period.

3.A.8 Proof of Theorem 3.2

Third period

We denote $x_{1,2} = \max\{x_1, x_2\}$. The maximum value of the problem is a piecewise function, depending on whether $x_{1,2}$ belongs to the learning region.

$$\mathcal{V}_3(x_{1,2}) = \begin{cases} \mu & \text{if } x_{1,2} \in [0, \mu - \frac{1}{4k}); \\ \\ kx_{1,2}^2 + ax_{1,2} + b & \text{if } x_{1,2} \in [\mu - \frac{1}{4k}, \mu + \frac{1}{4k}] \\ \\ x_{1,2} & \text{if } x_{1,2} \in (\mu + \frac{1}{4k}, 1]; \end{cases}$$

Function $\mathcal{V}_3(x_{1,2})$ is basically a value function of the static problem with outside option $x_{1,2}$.

Second period

We consider several cases with different values of x_1.

1. Case $x_1 < \mu - \frac{1}{4k}$

In this case $\mathcal{V}_3(x_{1,2}) = \mathcal{V}_3(x_2)$ simply by the construction. Therefore, $u_2(x_1, x_2) = \mathcal{V}_3(x_2) - k(x_2 - \mu)^2$. The solution is identical to the two-period problem, thus $\mathcal{V}_2(x_1) = \mu + \frac{1}{16k}$

2. Case $x_1 > \mu + \frac{1}{4k}$ In this case

$$\mathcal{V}_3(x_{1,2}) = \begin{cases} x_1 & \text{if } x_1 > x_2; \\ \\ x_2 & \text{if } x_2 \geq x_1 \end{cases}$$

Therefore

$$u_2(x_1, x_2) = \begin{cases} x_1 - k(x_2 - \mu)^2 & \text{if } x_1 > x_2; \\ \\ x_2 - k(x_2 - \mu)^2 & \text{if } x_2 \geq x_1 \end{cases}$$

In this case, the problem is static with outside option x_1. Because $x_1 > \mu + \frac{1}{4k}$ the manager simply chooses the first candidate and $\mathcal{V}_2(x_1) = x_1$.

3. Case $x_1 \in [\mu - \frac{1}{4k}, \mu + \frac{1}{4k}]$

We denote $g(x) = kx^2 + ax + b$. Therefore, in this case

$$\mathcal{V}_3(x_{1,2}) = \begin{cases} g(x_1) & \text{if } x_2 < x_1; \\ \\ g(x_2) & \text{if } x_2 \in [x_1, \mu + \frac{1}{4k}] \\ \\ x_2 & \text{if } x_2 \in (\mu + \frac{1}{4k}, 1]; \end{cases}$$

Therefore

$$u_2(x_1, x_2) = \begin{cases} g(x_1) - k(x_2 - \mu)^2 & \text{if } x_2 < x_1; \\ \\ g(x_2) - k(x_2 - \mu)^2 & \text{if } x_2 \in [x_1, \mu + \frac{1}{4k}] \\ \\ x_2 - k(x_2 - \mu)^2 & \text{if } x_2 \in (\mu + \frac{1}{4k}, 1]; \end{cases}$$

We claim that the optimal posterior distribution p_2 has two posterior beliefs $x_2^L =$

$g(x_1) - \frac{1}{4k}$ and $x_2^H = g(x_1) + \frac{1}{4k}$.

Posterior distribution p_2 from above is the solution to static problem (3.1) with outside option $g(x_1)$. Therefore the concave closure for the static problem with with outside option $g(x_1)$ equals function $u_2(x_1, x_2)$ if $x_2 \in [0, x_2^L)$ and $x_2 \in (x_2^H, 1]$ and is the straight line that connects points $(x_2^L, u_2(x_1, x_2^L))$ and $(x_2^H, u_2(x_1, x_2^H))$. We denote this concave closure as $w_2(x_2)$. We show that function $w_2(x_2)$ also is a concave closure in our problem.

Function $w_2(x_2)$ is concave by construction. Inequality $w_2(x_2) \geq u_2(x_1, x_2)$ holds. Simply by the construction $w_2(x_2) = u_2(x_1, x_2)$ if $x_2 \in [0, x_2^L]$ and $x_2 \in [x_2^H, 1]$. Therefore $w_2(x_2) > u_2(x_1, x_2)$ if $x_2 \in (x_2^L, x_1]$ and $x_2 \in (\mu + \frac{1}{4k}, x_2^H]$ hold because the concave function lies below the tangent line. The line, that connects points $(x_1, u_2(x_1, x_1))$, $(\mu + \frac{1}{4k}, u_2(x_1, \mu + \frac{1}{4k}))$ is tangent to function $u_2(x_1, x_2)$ at point $x_2 = \mu + \frac{1}{4k}$; the line, that connects points $(x_2^L, u_2(x_1, x_2^L))$, $(x_2^H, u_2(x_1, x_2^H))$ is tangent to function $u_2(x_1, x_2)$ at point $x_2 = x_2^H$. Because function $u_2(x_1, x_2)$ is concave in x_2 on $x_2 \in (\mu + \frac{1}{4k}, x_2^H)$ inequality $w(x_2) > u_2(x_1, x_2)$ holds on $x_2 \in (x_1, \mu + \frac{1}{4k})$ also holds.

We assume that concave function $\tilde{w}_2(x_2)$ exists such that $\tilde{w}_2(x_2) \geq u_1(x_1, x_2)$ for all x_2 and point $\tilde{x}_2$ exists such that $w_2(\tilde{x}_2) > \tilde{w}_2(\tilde{x}_2)$. Clearly $\tilde{x}_2 \in (x_2^L, x_2^H)$ holds. Consider the line $\tilde{f}(x_2) = \tilde{\alpha} x_2 + \tilde{\beta}$ such that it equals $\tilde{w}_2(x_2)$ at $x_2 = \tilde{x}_2$ and is larger than $\tilde{w}_2(x_2)$ if $x_2 \neq \tilde{x}_2$. Consider also the line that is tangent to $w_2(x_2)$ at $x_2 = \tilde{x}_2$. We denote the line as $f(x) = \alpha x_2 + \beta$. Because $\tilde{f}(x_2^H) \geq f(x_2^H)$ inequality $\tilde{\alpha} > \alpha$ holds. Additionally, because $\tilde{f}(x_2^L) \geq f(x_2^H)$ inequality $\tilde{\alpha} < \alpha$ also holds, which leads to the contradiction, and therefore function $w_2(x_2)$ also is a concave closure in our problem.

In this case equality $\mathcal{V}_2(x_1) = g(g(x_1))$ holds, therefore, the value of the problem is the same because it is the same as in the static problem with outside option $g(x_1)$.

First period

From the analysis above, the net utility in the first period equals

$$u_1(x_1) = \begin{cases} \mu + \frac{1}{16k} - k(x_1-\mu)^2 & \text{if } x_1 < \mu - \frac{1}{4k}; \\ \\ g(g(x_1)) - k(x_1-\mu)^2 & \text{if } x_1 \in [\mu - \frac{1}{4k}, \mu + \frac{1}{4k}] \\ \\ x_1 - k(x_1-\mu)^2 & \text{if } x_2 \in (\mu + \frac{1}{4k}, 1]; \end{cases}$$

We claim that the optimal distribution p_1 has two posterior beliefs x_1^L, x_1^H such that $x_1^L \in (\mu - \frac{1}{4k}, \mu)$ and $x_1^H \in (\mu + \frac{1}{4k}, 1)$. Moreover the straight line that connects points $(x_1^L, u(x_1^L))$ and $(x_1^H, u(x_1^H))$ is tangent to function $u_1(x_1)$ at points $x_1 = x_1^L$ and $x_1 = x_1^H$.We consider function $w_1(x_1)$ which equals $u_1(x_1)$ if $x_1 \notin (x_1^L, x_1^H)$ and equals to the straight line that connects points $(x_1^L, u(x_1^L))$ and $(x_1^H, u(x_1^H))$. We show that $w_1(x_1)$ is a concave closure of $u_1(x_1)$.

We first show that function $w_1(x_1)$ exists. To show this, we need to show that there is a common tangent line to function $u_1(x_1)$ on the intervals $(\mu - \frac{1}{4k}, \mu)$ and $(\mu + \frac{1}{4k}, 1)$.

To analyze tangent lines to function $u_1(x_1)$, we find the behavior of the first and the second derivative of the function.

Function $u_1(x_1)$ is continuously differentiable because simple algebra shows that $[g(g(\mu - \frac{1}{4k}))]' = 0$ and $[g(g(\mu + \frac{1}{4k}))]' = 1$. Simple algebra also shows that the second derivative of $g(g(x_1))$ is

$$[g(g(x_1))]'' = 12k(k^2(x-\mu)^2 + \frac{k}{2}(x-\mu) + \frac{7}{48})$$

Therefore $u''(x_1) = 12k(k^2(x-\mu)^2 + \frac{k}{2}(x-\mu)) - \frac{1}{4}k.$ on interval $[\mu - \frac{1}{4k}, \mu + \frac{1}{4k}]$.

The minimum of the second derivative is at point $x_1 = \mu - \frac{1}{4k}$. Because of inequalities $u''(\mu) < 0$ and $u''(\mu + \frac{1}{4k})) > 0$ the second derivative has a root $\hat{x}_1$ on the interval $(\mu, \mu + \frac{1}{4k})$. Therefore, $g''(g(x_1)) < 0$ is negative on $(\mu - \frac{1}{4k}, \hat{x}_1)$ and positive on $(\hat{x}_1, \mu + \frac{1}{4k})$. Thus, $u_1'(x_1)$ decreases and $u_1(x_1)$ is concave on $(\mu - \frac{1}{4k}, \hat{x}_1)$; $u_1'(x_1)$ increases and $u_1(x_1)$ is convex on $(\hat{x}_1, \mu + \frac{1}{4k})$.

We observe that $u'(\mu - \frac{1}{4}) = u'(\mu + \frac{1}{4}) = \frac{1}{2}$. Additionally, $u'(\mu) = \frac{5}{16}$ holds and point $\tilde{x}_1 = \mu + \frac{11}{32k} \in (\mu + \frac{1}{4k}, 1)$ exists, such that $u'(\tilde{x}_1) = \frac{5}{16}$. Because function $u'(x_1)$ is decreasing on $(\mu - \frac{1}{4k}, \mu)$ this function is a one-to-one function on this interval. Every value of $u'(x_1)$ is associated with the intercept of the tangent line to function $u_1(x_1)$ at point x_1. We denote the slope of the tangent line to $u(x_1)$ on interval $[\mu - \frac{1}{4k}, \mu]$ as α. Because $u'(x_1)$ is a one-to-one function on $x_1 \in [\mu - \frac{1}{4k}, \mu]$, an implicit mapping $\beta(\alpha)$ that defines the value of the intercept given the α is a function. Moreover, because function $u_1(x_1)$ is continuously differentiable, function $\beta(\alpha)$ is continuous. Similarly, we denote the slope of the tangent line to $u(x_1)$ on interval $[\mu + \frac{1}{4k}, \tilde{x}_1]$ as $\tilde{\alpha}$. Analogously, implicit mapping $\tilde{\beta}(\tilde{\alpha})$ is also a continuous function.

Because $u'(\mu - \frac{1}{4}) = u'(\mu + \frac{1}{4})$ and $u'(\mu) = u'(\tilde{x}_1)$ hold, slopes β and $\tilde{\beta}$ take the same values on intervals $x_1 \in [\mu - \frac{1}{4}, \mu]$ and $x_1 \in [\mu + \frac{1}{4}, \tilde{x}_1]$. Therefore, we can consider a function $\bar{\beta}(\bar{\alpha}) = \beta(\bar{\alpha}) - \tilde{\beta}(\bar{\alpha})$, where $\bar{\alpha} \in [\frac{5}{16}, \frac{1}{2}]$. Function $\bar{\beta}(\bar{\alpha})$ is continuous because functions $\beta(\alpha)$ and $\tilde{\beta}(\tilde{\alpha})$ are continuous. We observe that $\bar{\beta}(\frac{1}{2}) = \frac{1}{16k} > 0$. We claim that $\bar{\beta}(\frac{5}{16}) < 0$. The intercept at point $x_1 = \tilde{x}_1$ equals $\tilde{\beta}(\frac{5}{16}) = \frac{11}{16}\mu + \frac{121}{1024k}$. The value of the tangent line at point $\tilde{x}_1$ if $x = \mu$ equals $\mu + \frac{121}{1024k}$. Because of the equality $g(g(\mu)) = \mu + \frac{100}{1024k}$ the tangent line at points $\tilde{x}_1$ lies above function $u_1(x_1)$ at point $x_1 = \mu$. Because $u_1'(\mu) = u_1'(\tilde{x}_1)$ inequality $\beta(\mu) < \tilde{\beta}(\frac{5}{16})$ holds and therefore $\bar{\beta}(\frac{5}{16}) < 0$.

Function $\bar{\beta}$ is continuous on $[\frac{5}{16}, \frac{1}{2}]$ and inequality $\bar{\beta}(\frac{1}{2})\bar{\beta}(\frac{5}{16}) < 0$ holds. Therefore, point $\alpha' \in (\frac{5}{16}, \frac{1}{2})$ exists, such that $\bar{\beta}(\alpha') = 0$. The line $y = \alpha' x + \beta(\alpha')$ is the common tangent line to function $u_1(x_1)$ on intervals $(\mu - \frac{1}{4k}, \mu)$ and $(\mu + \frac{1}{4k}, 1)$. We denote x_1 such that $u_1'(x_1) = \alpha'$ on interval $(\mu - \frac{1}{4k}, \mu)$ as x_1^L, and on interval $(\mu + \frac{1}{4k}, 1)$ as x_1^H and therefore function $w_1(x_1)$ exists.

Function $w_1(x_1)$ is concave by construction. Because of inequality $x_1^L < \mu$, inequality $x_1^L < \hat{x}_1$ holds. Inequality $w_1(x_1) \geq u_1(x_1)$ holds on intervals $[x_1^L, \hat{x}_1]$ and $[\mu + \frac{1}{4k}, x_1^H]$ because function $u_1(x_1)$ is concave on them. Function $w_1(x_1)$ is larger than function $u_1(x_1)$ on the boundaries of interval $(\hat{x}_1, \mu + \frac{1}{4k})$. Because function $u_1(\hat{x}_1)$ is convex on this interval, inequality $w_1(x_1) > u_1(\hat{x}_1)$ holds for all $x_1 \in (\hat{x}_1, \mu + \frac{1}{4k})$. Therefore, inequality $w_1(x_1) \geq u_1(\hat{x}_1)$ holds for all $x_1 \in [0, 1]$. The proof that function $w_1(x_1)$ is the minimal concave function is identical to the proof for the second period above for function $w_2(x_2)$. Therefore $w_1(x_1)$ is the concave closure of $u_1(x_1)$.

Solution to the problem

From the analysis above in the solution to the problem the manager chooses p_1 with two posteriors $x_1^L \in (\mu - \frac{1}{4k}, \mu), x_1^H \in (\mu + \frac{1}{4k}, 1)$. After high posterior realization, the manager does not consume informative signals about candidates in later stages and chooses the first candidate. After low posterior realization, the manager chooses p_2 with two posteriors

$x_2^L = g(x_1^L) - \frac{1}{4k}, x_2^H = g(x_1^L) + \frac{1}{4k}$. Because function $g(x)$ increases when $x > \mu - \frac{1}{4k}$ inequalities $g(\mu - \frac{1}{4k}) < g(x_1^L) < g(\mu)$ hold. Because $g(\mu - \frac{1}{4k}) = \mu, g(\mu) = \mu + \frac{1}{16k}$ inequalities

$$x_2^L \in (\mu - \frac{1}{4k}, \mu - \frac{3}{16k}); \quad x_2^H \in (\mu + \frac{1}{4k}, \mu + \frac{5}{16k})$$

hold. Therefore, after realization x_2^H the manager does not consume informative signals about the third candidate and chooses the second candidate.

Simple algebra shows that equality $g(x) - x - \frac{1}{4k} = k(x - (\mu + \frac{3}{4k}))(x - (\mu - \frac{1}{4k}))$ holds. Therefore, on interval $x \in (\mu - \frac{1}{4k}, \mu)$ inequality $x > g(x) - \frac{1}{4k}$ holds. Thus, $x_1^L > x_2^L$.

In the third period the manager chooses p_3 with two posteriors $x_3^L = x_1^L - \frac{1}{4k}, x_3^H = x_1^L + \frac{1}{4k}$. After high posterior realization the manager chooses the third candidate. After low posterior realization the manager chooses the first candidate.

By the construction of function $g(x)$ inequality $g(x) > x$ holds if $x \in (\mu - \frac{1}{4k}, \mu + \frac{1}{4k})$. Therefore inequalities $x_3^L < x_2^L, x_3^H < x_2^H$ hold. We claim that inequality $x_1^H > x_2^H$ holds. We compare the slopes of the straight lines that belong to functions $w_1(x_1)$ and $w_2(x_2)$. We show that the slope of a line that is tangent to $w_2(x_2)$ is greater than a line tangent to $w_1(x_1)$. Because $w_1(x) = w_2(x)$ if $x \in (\mu + \frac{1}{4k}, 1)$ and function $w_1(x)$ is concave on this interval, inequalities of the slopes imply that $x_1^H > x_2^H$.

The slope of the straight line that belongs to function $w_1(x_1)$ equals $g'(g(x_1^L)g'(x_1^L)$; the slope of the straight lines, that belong to function $w_2(x_2)$ equals $-2k(x_2^L - \mu)$. Simple algebra shows that $g'(g(x)g'(x) + 2k(g(x) - \frac{1}{4k} - \mu) = 4k^3(x - (\mu - \frac{3}{4k}))(x - (\mu - \frac{1}{4k}))(x - (\mu + \frac{1}{4k}))$. Therefore, inequality $g'(g(x)g'(x) < -2k(x - \mu)$ holds if $x \in (\mu - \frac{1}{4k}, \mu + \frac{1}{4k})$. Thus, inequality $g'(g(x_1^L)g'(x_1^L) < -2k(x_1^L - \mu)$ holds and $x_1^H > x_2^H$. From the analysis above the linear order on the optimal posterior beliefs is

$$x_3^L < x_2^L < x_1^L < x_3^H < x_2^H < x_1^H.$$

3.A.9 Proof of Corollary 3.2

The manager can choose the second or the third candidate only if she receives the low signal in the first period. This event enters into the formula for both unconditional probabilities of choice, making it redundant for comparison. In later calculations, we omit the probability of the event. By the Bayes rule, the unconditional probability of choosing the second candidate equals

$$P(\text{choose } 2) = \frac{1}{2} + 2k(\mu - g(x_1^L)); \quad \text{P(choose 3)} = \Big(\frac{1}{2} - 2k(\mu - g(x_1^L))\Big)\Big(\frac{1}{2} + 2k(\mu - x_1^L)\Big).$$

Simple algebra shows that $P(\text{choose } 2) - P(\text{choose } 3) = 4k^3(x - (\mu + \frac{1}{4k}))^2(x - (\mu - \frac{1}{4k}))$. Because $x_1^L > \mu - \frac{1}{4k}$ inequality $P(\text{choose } 2) > P(\text{choose } 3)$ holds.

By the Bayes rule and the proof of Corollary 3.1, the average learning costs in the second and the third periods equal

$$C_2(p_2) = -k\Big(g(x_1^L) - \mu\Big)^2 + \frac{1}{16k}; \quad C_3(p_3) = \Big(\frac{1}{2} - 2k(\mu - g(x_1^L))\Big)\Big(-k\Big(x_1^L - \mu\Big)^2 + \frac{1}{16k}\Big)$$

Simple algebra shows that $C_2(p_2) - C_3(p_3) = k^3\Big(x - (\mu + \frac{1}{4k})\Big)^2\Big((x - (\mu - \frac{1}{4k}))^2 + \frac{1}{4k^2}\Big)$. Therefore, $C_2(p_2) > C_3(p_3)$.

Because inequality $g(x_1^L) > x_1^L$ holds, the amount of information that the manager obtains in the third period, conditional on the manager being in the period and choosing informative signals in this period, is larger than the amount she obtains in the second period.

3.A.10 Proof of Corollary 3.5

To characterize the solution to the problem, we apply arguments from the proof of Theorem 3.2. Because expected qualities are different $\mu_2 \neq \mu_3$ it is convenient to introduce function $g(x, \mu) = kx^2 + a_\mu x + b_\mu$. With outside option $x_1 \in (\mu_3 - \frac{1}{4k}, \mu_3 + \frac{1}{4k})$ net utility in the second period equals

$$u_2(x_1, x_2) = \begin{cases} g(x_1, \mu_3) - k(x_2 - \mu_2)^2 & \text{if } x_2 < x_1; \\ g(x_2, \mu_3) - k(x_2 - \mu_2)^2 & \text{if } x_2 \in [x_1, \mu_3 + \frac{1}{4k}] \\ x_2 - k(x_2 - \mu_2)^2 & \text{if } x_2 \in (\mu_3 + \frac{1}{4k}, 1]; \end{cases}$$

From the proof of Theorem 3.2 distribution p_2 with posterior beliefs $x_2^L = g(x_1, \mu_3) - \frac{1}{4k}, x_2^H = g(x_1, \mu_3) + \frac{1}{4k}$ is optimal if inclusion $\mu_2 \in (x_2^L, x_2^H)$ holds.

Because of inequality $g(x_1, \mu_3) < \mu_3$ for inequality $\mu_2 < g(x_1, \mu_3) + \frac{1}{4k}$ to hold it is sufficient to have $\mu_2 < \mu_3 + \frac{1}{4k}$. Because of inequality $g(x_1, \mu_3) > \mu_3 - \frac{1}{4k}$ exists δ_μ such that if $\mu_2 > \mu_3 - \delta_\mu$ inequality $\mu_2 > g(x_1, \mu_3) - \frac{1}{4k}$ holds. Therefore, if $\mu_2 \in (\mu_3 - \delta_\mu, \mu_3 + \frac{1}{4k})$ distribution p_2 defined above is optimal in the second stage.

From the analysis of the static problem in the third stage, the manager consumes an information signal p_3 such that $x_3^L = x_1 - \frac{1}{4k}, x_3^H = x_1 + \frac{1}{4k}$ only in the case of receiving posterior x_2^L.

We move to the optimal order analysis. With the notation above, the value of the problem equals $g(g(\mu_3, x_1), \mu_2)$. We show that inequality $\frac{dg(g(\mu_3,x_1),\mu_2)}{d\mu_2}\Big|_{\mu_2=\mu_3} > \frac{dg(g(\mu_3,x_1),\mu_2)}{d\mu_3}\Big|_{\mu_2=\mu_3}$ holds. If candidate i is marginally better than candidate j, the manager prefers to inspect candidate i if the last inequality holds.

The derivatives can be expressed as

$$\frac{dg(g(\mu_3,x_1),\mu_2)}{d\mu_2}\Big|_{\mu_2=\mu_3} = \frac{1}{2} = 2k(g(x_1,\mu) - \mu)$$
$$\frac{dg(g(\mu_3,x_1),\mu_2)}{d\mu_3}\Big|_{\mu_2=\mu_3} = (\frac{1}{2} - 2k(\mu - g(x_1,\mu))(\frac{1}{2} - 2k(x_1 - \mu)),$$

where $\mu = \mu_2 = \mu_3$. Simple algebra shows that

$$\frac{dg(g(\mu_3,x_1),\mu_2)}{d\mu_2}\Big|_{\mu_2=\mu_3} > \frac{dg(g(\mu_3,x_1),\mu_2)}{d\mu_3}\Big|_{\mu_2=\mu_3} = 4k^3(x_1 - (\mu + \frac{1}{4k}))^2(x_1 - (\mu - \frac{1}{4k})).$$

Because $x_1 > \mu - \frac{1}{4k}$ inequality $\frac{dg(g(\mu_3,x_1),\mu_2)}{d\mu_2}\Big|_{\mu_2=\mu_3} > \frac{dg(g(\mu_3,x_1),\mu_2)}{d\mu_3}\Big|_{\mu_2=\mu_3}$ holds.

3.A.11 Proof of Corollary 3.7

The proof is almost identical to the proof of Corollary 3.5. Distribution p_2 with posterior beliefs $x_2^L = g(x_1) - \frac{1}{4k_2}, x_2^H = g(x_1) + \frac{1}{4k_2}$ is optimal if inclusion $\mu \in (x_2^L, x_2^H)$ holds. Because inclusion $\mu \in (g(x_1) - \frac{1}{4k_3}, g(x_1) + \frac{1}{4k_3})$ holds, $\delta_k > 1$ exists such that, if $k_2 < \delta_k k_3$ then $\mu \in (x_2^L, x_2^H)$ holds.

From the proof of Corollary 3.2 probabilities of choosing the second and the third candidates in the case of different marginal cost of information if $x_1 = \mu$ equal

$$P(\text{choose } 2) = \frac{1}{2} - \frac{k_2}{8k_3}; \quad \text{P(choose 3)} = \frac{1}{2}\Big(\frac{1}{2} + \frac{k_2}{8k_3}\Big).$$

Therefore, inequality $P(\text{choose } 2) > P(\text{choose } 3)$ holds iff $k_3 > \frac{3}{4}k_2$.

Also, from the proof of Corollary 3.2 the average information cost in the second and the third periods in the case of different marginal cost of information if $x_1 = \mu$ equal

$$C_2(p_2) = \frac{1}{16k_2} - \frac{k_2}{256k_3^2}; \quad C_3(p_3) = \frac{1}{32k_3} + \frac{k_2}{128k_3^2}$$

Simple algebra shows that inequality $C_2(p_2) > C_3(p_3)$ is equivalent to the inequality $k_3 > \frac{3}{4}k_2$.

Simple algebra shows that in the case $x_1 = \mu$ condition $\mu \in (x_2^L, x_2^H)$ simplifies to $k_3 > \frac{1}{4}k_2$. Therefore, if $k_3 \in (\frac{3}{4}k_2, k_2)$ then the manager incurs more cost in the second period and chooses the second candidate the most often. However, if $k_3 \in (\frac{1}{4}k_2, \frac{3}{4}k_2)$ then the manager incurs more cost in the third period and chooses the second candidate least often.

Bibliography

Aczél, J. and J. Pfanzagl (1967). Remarks on the measurement of subjective probability and information. *Metrika 11*(1), 91–105.

Aliprantis, C. D. and K. C. Border (2006). *Infinite Dimensional Analysis: A Hitchhiker's Guide* (Third ed.). Springer.

Alonso, R. and O. Câmara (2016). Bayesian persuasion with heterogeneous priors. *Journal of Economic Theory 165*, 672–706.

Alonso, R., W. Dessein, and N. Matouschek (2008). Centralization Versus Decentralization: An Application to Price Setting by a Multi-Market Firm. *Journal of the European Economic Association 6*(2-3), 457–467.

Alonso, R. and N. Matouschek (2008). Optimal delegation. *The Review of Economic Studies 75*(1), 259–293.

Argenziano, R., S. Severinov, and F. Squintani (2016). Strategic Information Acquisition and Transmission. *American Economic Journal: Microeconomics 8*(3), 119–155.

Armstrong, M. and D. E. Sappington (2007). Recent Developments in the Theory of Regulation. In *Handbook of Industrial Organization*, Volume 3, pp. 1557–1700.

Armstrong, M., J. Vickers, and J. Zhou (2009). Prominence and consumer search. *The RAND Journal of Economics 40*(2), 209–233.

Ball, I. and X. Gao (2021). Benefiting from bias. working paper.

Banerjee, A. and R. Somanathan (2001). A simple model of voice. *The Quarterly Journal of Economics 116*(1), 189–227.

Barberis, N. C. (2013). Thirty Years of Prospect Theory in Economics: A Review and Assessment. *Journal of Economic Perspectives 27*(1), 173–196.

Bartoš, V., M. Bauer, J. Chytilová, and F. Matějka (2016). Attention discrimination: Theory and field experiments with monitoring information acquisition. *American Economic Review 106*(6), 1437–75.

Bernheim, B. D. (1994). A theory of conformity. *Journal of Political Economy 102*(5), 841–877.

Bernheim, B. D. and C. Exley (2015). Understanding conformity: An experimental investigation. *Harvard Business School NOM Unit Working Paper* (16-070).

Caplin, A. and M. Dean (2013). Behavioral implications of rational inattention with Shannon entropy. Technical report, National Bureau of Economic Research.

Caplin, A., M. Dean, and J. Leahy (2019). Rational inattention, optimal consideration sets, and stochastic choice. *Review of Economic Studies 86*(3), 1061–1094.

Caplin, A., M. Dean, and J. Leahy (2022). Rationally inattentive behavior: Characterizing and generalizing shannon entropy. *Journal of Political Economy 130*(6), 1676–1715.

Che, Y.-K., W. Dessein, and N. Kartik (2013). Pandering to Persuade. *The American Economic Review 103*(1), 47–79.

Che, Y.-K. and N. Kartik (2009). Opinions as incentives. *Journal of Political Economy 117*(5), 815–860.

Cover, T. M. and J. A. Thomas (2012). *Elements of information theory.* John Wiley & Sons.

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

Demski, J. S. and D. E. Sappington (1987). Delegated expertise. *Journal of Accounting Research*, 68–89. Publisher: JSTOR.

Dessein, W. (2002). Authority and communication in organizations. *The Review of Economic Studies 69*(4), 811–838.

Dogan, M. and J. Hu (2022). Consumer search and optimal information. *The RAND Journal of Economics*.

Egorov, G. and K. Sonin (2011). Dictators and their viziers: Endogenizing the loyalty–competence trade-off. *Journal of the European Economic Association 9*(5), 903–930.

Ekinci, E. and N. Theodoropoulos (2021). Determinants of delegation: Evidence from british establishment data. *Bogazici Journal: Review of Social, Economic & Administrative Studies 35*(1).

Feigenbaum, E. A. and H. A. Simon (1962). A theory of the serial position effect. *British Journal of Psychology 53*(3), 307–320.

Fershtman, D. and A. Pavan (2022). Searching for arms: Experimentation with endogenous consideration sets. Technical report, Mimeo.

Flynn, J. P. and K. Sastry (2020). Strategic mistakes in large games. *Available at SSRN*.

Foss, N. J. and K. Laursen (2005). Performance pay, delegation and multitasking under uncertainty and innovativeness: An empirical investigation. *Journal of Economic Behavior & Organization 58*(2), 246–276.

Gilbert, J. P. and F. Mosteller (1966). Recognizing the maximum of a sequence. *Journal of the American Statistical Association 61*(313), 35–73.

Gossner, O., J. Steiner, and C. Stewart (2021). Attention please! *Econometrica 89*(4), 1717–1751.

Graham, J. R., C. R. Harvey, and M. Puri (2015). Capital allocation and delegation of decision-making authority within firms. *Journal of financial economics 115*(3), 449–470.

Hébert, B. and M. Woodford (2019). Rational Inattention when Decisions Take Time. Technical Report w26415, National Bureau of Economic Research, Cambridge, MA.

Hoffman, M., L. B. Kahn, and D. Li (2018). Discretion in hiring. *The Quarterly Journal of Economics 133*(2), 765–800.

Holmström, B. (1980). On the theory of delegation. Discussion Papers 438, Northwestern University, Center for Mathematical Studies in Economics and Management Science.

Jain, V. and M. Whitmeyer (2021). Search and competition with flexible investigations.

Jerath, K. and Q. Ren (2021). Consumer rational (in) attention to favorable and unfavorable product information, and firm information design. *Journal of Marketing Research 58*(2), 343–362.

Kamenica, E. (2017). Information economics. *Journal of Political Economy 125*(6), 1885–1890.

Kennedy, J. B. (2016). Who do you trust? presidential delegation in executive orders. *Research & Politics 3*(1).

Laffont, J.-J. and D. Martimort (2009). *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press.

Laffont, J.-J. and J. Triole (1990). The Politics of Government Decision Making: Regulatory Institutions. *Journal of Law, Economics, and Organization 6*(1), 1–32.

Li, H. and W. Suen (2004). Delegating decisions to experts. *Journal of Political Economy 112*(S1), S311–S335.

Li, Y. and N. Epley (2009). When the best appears to be saved for last: Serial position effects on choice. *Journal of Behavioral Decision Making 22*(4), 378–389.

Lindbeck, A., S. Nyberg, and J. W. Weibull (1999). Social norms and economic incentives in the welfare state. *Quarterly Journal of Economics 114*(1), 1–35.

Lindbeck, A. and J. Weibull (2020). Delegation of investment decisions, and optimal remuneration of agents. *European Economic Review 129*, 103559.

Lipnowski, E., L. Mathevet, and D. Wei (2020). Attention management. *American Economic Review: Insights 2*(1), 17–32.

Lipnowski, E., L. Mathevet, and D. Wei (2022). Optimal attention management: A tractable framework. *Games and Economic Behavior 133*, 170–180.

Liu, L., X. H. Wang, and H. Yu (2022). Sequential search with partial depth. *Economics Letters 216*, 110624.

Matějka, F. and A. McKay (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review 105*(1), 272–98.

Matveenko, A. and S. Mikhalishchev (2021). Attentional role of quota implementation. *Journal of Economic Theory 198*, 105356.

Mensch, J. (2018). Cardinal representations of information. *Available at SSRN 3148954*.

Miao, J. and H. Xing (2020). Dynamic discrete choice under rational inattention.

Morris, S. (1995). The common prior assumption in economic theory. *Economics & Philosophy 11*(2), 227–253.

Newell, B. R. and D. R. Shanks (2014). Unconscious influences on decision making: A critical review. *Behavioral and brain sciences 37*(1), 1–19.

Nimark, K. P. and S. Sundaresan (2019). Inattention and belief polarization. *Journal of Economic Theory 180*, 203–228.

Prendergast, C. (1993). A theory of "Yes Men". *The American Economic Review 83*(4), 757–770.

Rantakari, H. (2008). Governing Adaptation. *The Review of Economic Studies 75*(4), 1257–1285.

Rényi, A. (1961). On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, pp. 547–561. University of California Press.

Savage, L. J. (1971). Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association 66*(336), 783–801.

Shore, L. M., B. G. Chung-Herrera, M. A. Dean, K. H. Ehrhart, D. I. Jung, A. E. Randel, and G. Singh (2009). Diversity in organizations: Where are we now and where are we going? *Human resource management review 19*(2), 117–133.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics 69*(1), 99–118.

Simon, H. A. et al. (1971). Designing organizations for an information-rich world. *Computers, communications, and the public interest 72*, 37.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665–690.

Steiner, J. and C. Stewart (2016). Perceiving Prospects Properly. *American Economic Review 106*(7), 1601–1631.

Steiner, J., C. Stewart, and F. Matějka (2017). Rational inattention dynamics: Inertia and delay in decision-making. *Econometrica 85*(2), 521–553.

Szalay, D. (2005). The economics of clear advice and extreme options. *The Review of Economic Studies 72*(4), 1173–1198.

Tsakas, E. (2020). Robust scoring rules. *Theoretical Economics 15*(3), 955–987.

Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty 5*(4), 297–323.

Wei, D. (2021). Persuasion under costly learning. *Journal of Mathematical Economics 94*, 102451.

Weitzman, M. L. (1979). Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*, 641–654.