Stealth Startups, Clauses, and Add-ons: A Model of Strategic Obfuscation

Monte, Daniel and Linhares, Luis Henrique

9 January 2023

Stealth Startups, Clauses, and Add-ons: A Model of Strategic Obfuscation

Luis Henrique Linhares Daniel Monte*

January 9, 2023

Abstract

Obfuscation is ubiquitous and often intentional. We consider an uninformed Principal who chooses how costly it will be for the Agent to obtain and process new information. Thus, obfuscation and transparency are endogenous to the problem at hand. Using a rational inattention framework, we study the Principal's optimal induced cost of processing information and examine necessary and sufficient conditions for obfuscation. We characterize the Principal's optimal obfuscation for the class of state independent preferences. We apply our model to examples such as stealth startups, companies with unnecessarily complicated contracts, and firms whose products have varying features that disguise add-ons.

Keywords: stealth startups, information design, rational inattention

JEL Codes: D11, D82, D83

*Linhares: Boston University. E-mail: lhrlinha@bu.edu. Monte: ESOMAS department, University of Turin, daniel.monte@unito.it and Sao Paulo School of Economics - FGV. E-mail: daniel.monte@fgv.br. We are very thankful for conversations with Henrique de Oliveira, Lucas Maestri, Chiara Margaria, Laurent Mathevet, Juan Ortner, Kareen Rozen, and Adrien Vigier.

1 Introduction

A well-known phenomenon in the venture capital community is that of stealth startups. These are companies that, before launching, operate in stealth mode, a temporary state of secrecy. Companies may wish to keep their ideas or products secret and attract investment before they know their flagship product's full potential. An example is CNEX Lab, a private semiconductor company founded in 2013 that raised over $38 million in under five years while in stealth mode.[1] Their website had a single line: "Leading the evolution in big data storage." The most common justification for operating in stealth mode is that founders believe their new technology to be disruptive. The downside is that operating in secrecy conceals valuable information about the product and its potential profitability.[2] Should a founder still unsure about her company's profitability keep her company in stealth mode? Should she reveal partial information, or, perhaps, as much information as possible?

Indeed, obfuscation, by which we mean the conscious effort to hinder the collection and processing of information, is ubiquitous. When required to disclose evidence, lawyers may send excess material not directly linked to the case, making it harder for other parties to discern what is relevant and what is not. Firms often disguise add-ons: for example, printers can be sold at a loss to attract consumers while hiding ink sales at high margins.

Empirically, there is abundant evidence of obfuscation efforts. For example, competing firms regularly charge different prices for similar products, as obfuscation can artificially create perceived product-differentiation in the eyes of the consumer. The mechanisms are numerous: by offering various product sizes and flavors (Richards et al., 2019), by making it more difficult to compare prices online (Ellison and Ellinson, 2009), or, as previously mentioned, by hiding the price of add-ons (Ellison, 2005). Other studies lay out how obfuscation may be sustained in a more competitive market: Gabaix and Laibson (2006) illustrate cases in

[1] https://www.cbinsights.com/research/well-funded-stealth-tech-startups/

[2] Indeed, many stealth startups do not live up to expectations; Segway Transporter is perhaps the most well-known example. See https://www.economist.com/technology-quarterly/2010/06/12/mr-segways-difficult-path

which firms do not have an incentive to educate consumers about their competitors, confusing pricing can make comparing difficult for consumers (Piccione and Spiegler, 2012), and Jin et al. (2018) show that even companies with high-quality products and services may try to confuse consumers as this induces systematic mistakes that they may exploit.

In our model, a Principal wishing to persuade an Agent chooses **how costly** it will be for the Agent to process information. We model this cost of processing information as in the Rational Inattention literature, detailed in the seminal paper by Sims (2003). The idea of Rational Inattention is that the Agent is rational and information is available, but obtaining or processing it comes at a cost. So, when choosing how much to learn, the Agent will weigh the costs and benefits of finding out the true state of the world. To capture costly information processing, we use the expected reduction of Shannon Entropy (Shannon, 1948).

The Principal wishes to induce the Agent to take a particular action. Before choosing an action, the Agent will perform their investigations to try and acquire further information in the form of a signal whose realization is conditioned on the state of the world. The caveat is that the Agent is rationally inattentive and has to spend time and effort to obtain and process information. We assume that the Principal can facilitate or hinder these efforts by choosing the information processing cost level. Given the costs, the Agent will choose an optimal experiment and take action based on her induced beliefs. More formally, the Principal can choose the value of a parameter $k \in \mathbb{R}$ that linearly affects how much uncertainty (measured by Shannon Entropy) is reduced by the acquired information.[3] Thus, by setting $k = 0$, the Principal is effectively choosing to let the Agent correctly identify the state of the world. By setting a large k, think of $k \to \infty$, the Principal is making any new information processing impossibly costly so that the Agent will base her decisions on the common prior.

The Principal's efforts to raise costs, or increase the value of k, can be interpreted in different ways: restricting access to information, decreasing the time available to make a decision, or simply dumping irrelevant data. Unlike other obfuscation papers, we do not focus

[3] This parameter is standard in the Rational Inattention literature that uses Shannon Entropy: $k(H(\mu_0) - \mathbb{E}_\tau[H(\mu_a)])$.

on a single mechanism (for instance, the time to make a decision). Instead, we generalize how this behavior can appear if the Agent displays rational inattention.

We consider that the value of k is chosen before the Principal finds out the state of the world (thus, the chosen k is not a signal in itself). We comment further on this assumption and why we consider it reasonable when presenting the model and describing practical examples.[4]

Consider the following example of investment in startups. An investor is deciding whether or not to invest in a startup. The startup's founder wishes to persuade the investor to invest in her company. The product may be a disruptive technology, implying that the startup will generate very high profits, or it may be a neutral technology, not worth investing in. The founder and the investor share a common prior $\varepsilon \in (0,1)$ that the technology is disruptive. The payoffs may be visualized below.

Table 1: Payoffs table for Startup Investment game.

		Investor	
		not invest	invest
Startup	Disruptive	$0,0$	$X,\ B$
	Neutral	$0,0$	$X,-C$

In which $X > 0$ and $B > C > 0$. Given these payoffs, there is a belief threshold $\varepsilon_{ind} = \frac{C}{B+C}$ above which the investor's choice is to invest. If this is the case, i.e., if the common belief that the product is disruptive is sufficiently high ($\varepsilon \geqslant \varepsilon_{ind}$), the startup will operate in complete stealth mode. If the belief is lower than the threshold, the startup must release information about her product to investors, or it will not receive investment. How much information will be optimal to release will depend on the payoffs under the different scenarios. For example, choosing to be completely transparent means that the investor will fully learn the underlying true state of the world, and there will only be investment if the product is disruptive. Since the prior is small, this will happen with a small probability. A more profitable policy for the startup might be to release information, but with high processing costs. By doing this, it will

[4] Martin (2017) and Clippel and Rozen (2020) present models in which there is a strategic component to the fact that an agent is choosing to obfuscate.

be too costly for the investor to learn the state of the world, and she will instead only gather partial information. After observing the outcome of the experiment in which only partial information is learned, the investor may have become more convinced that it is not worth investing, or it may have tilted her belief in a way that makes her sufficiently convinced that the investment is worthy. On average, this partial information, which leads to a small belief change, might be more profitable to the startup than a full-learning policy (and also more profitable than a policy of complete secrecy). With the techniques that we will present in our paper, we will be able to (i) compute exactly how costly the founder will choose information processing to be; (ii) understand how the investor chooses how much to learn given the chosen information processing costs; and (iii) perform some comparative statics given policies motivated at improving welfare, such as policies on consumer transparency. We discuss this example in detail in Section 5.

It is natural to compare our model with the Bayesian Persuasion framework.[5] There, a Sender commits to an information disclosure policy, while here, the Principal (essentially) chooses how the Agent will learn, within the rational inattention framework. A contribution of our paper is that it advances the idea that persuasion might be about selecting which information might be learned rather than about commitment in the information transmission stage.[6] As such, we have a story about a Principal trying to persuade an Agent without needing a strong commitment assumption.

In general, we provide a new framework in which, contrary to most of Bayesian Persuasion and the Rational Inattention literature, we endogenize how costly it will be for the Agent to obtain new information.

To better fix the idea behind the model, we provide applications that illustrate how the mechanism of obfuscation works, and through them, we also obtain some interesting insights.

[5]See Kamenica and Gentzkow (2011). See also Kamenica (2019) and Bergemann and Morris (2019) for comprehensive surveys.

[6]Indeed, Kamenica and Gentzkow (2011) use commitment in the information acquisition stage instead of commitment in the information transmission stage as a possible motivation for the commitment assumption in Bayesian Persuasion. Barros (2021) develops a model in which the Principal has commitment in the acquisition stage, but not in the transmission stage.

One is that there are two main factors that affect the optimal level of obfuscation: the preferences (if they are symmetric, or how aligned the agents are) and the prior.

Another interesting discussion is that obfuscation has clear welfare implications. When the Principal tries to obfuscate in order to increase expected payoffs, it may come at a high cost for the Agent. Therefore, understanding the reasons why an obfuscation behavior arises provides insights for policies that seek to increase the Agent's (who can be representing the consumers, for instance) welfare and provides valuable input for regulatory policies on mandated disclosure. Ben-Shahar and Schneider (2014), for example, detail how ineffective policies of mandated disclosure can be. While at first they look appealing from an economic perspective, (they tend not to distort the market as much as other policies) they do not bring about much change as firms engage in other types of obfuscation; it is common for customers to get lost with the excess of clauses in a contract, for instance. In one of our examples, we see that an effective policy - that at first glance seems to harm consumers -can induce firms to simplify disclosure as difficulty in processing information might push customers to not purchase their products.

We organize this paper as follows: Section 2 provides a brief review of the literature; Section 3 describes the theoretical model; Section 4 details our efforts to classify the problems by the level of obfuscation chosen; Section 5 provides practical applications of the model; Section 6 concludes.

2 Related Literature

Information Design and the commitment assumption - Our paper is related to the large literature in Bayesian Persuasion and information design (Kamenica and Gentzkow (2011), Bergemann and Morris (2019), for example). More directly related to our paper are the papers that relax the commitment assumption in information transmission in the Bayesian Persuasion models. For example, Lin and Liu (2022), Barros (2021), Lipnowski et al. (2022), Min (2021), Fréchette et al. (2022), Nguyen and Tan (2021), Perez-Richet and Skreta (2022) and Alonso and Camara (2022). Also related are Edmond and Lu (2017) who add confusion to the communication model by making the Sender capable of confusing the Agent at a cost and Bizzotto et al. (2020) who study the case of persuasion when dealing with certifications and add a cost for obtaining information.

Rational Inattention - This literature started with the seminal paper of Sims (2003), who used Shannon Entropy (Shannon, 1948) as it relates to a channel's capacity to send information. A branch of the literature developed mathematical foundations for the solutions of an agent that presents Rational Inattention: Matějka and McKay (2015) use Rational Inattention as a theoretic foundation for the use of the logit function to model discrete choices; in the same direction, there is a set of papers (Caplin and Dean, 2013, 2015; Caplin et al., 2022, 2019) that detail not only the solutions of the rational inattentive agent but also discuss some properties of using different cost functions. de Oliveira et al. (2017) characterizes an axiomatic foundation for the preferences of an agent that presents rational inattention, providing a way to generalize this class of problems. Ellis (2018) shows an axiomatic model of choice behavior for an agent with limited attention and gives the conditions in which the behavior can be considered the result of a choice of an optimal level of inattention. Martin (2017) uses Rational Inattention to explain how increasing attention costs can influence price strategies for companies trying to induce consumers to buy a product they do not know the intrinsic quality. These papers provide some of the mathematical foundations for the results we present.

Information Design + Rational Inattention - A few authors have recently combined the study of persuasion with Rational Inattention. Gentzkow and Kamenica (2014) is a direct extension of the classic Bayesian Persuasion model in which the sender has to buy the signal at a cost given by Shannon Entropy. Matyskova and Montes (2021) develop a model in which the Sender sends information at no cost. However, upon seeing a signal realization the Agent can decide if she wants to buy further information at a cost (given by the usual Shannon

Entropy). They derive how the sender will choose the optimal signal and a few properties that simplify the Sender's problem. Bloedel and Segal (2020) detail how to optimally persuade a rationally inattentive Agent. Knowing that a fully informative signal will not be processed and thus the concavication results from Kamenica and Gentzkow (2011) are not attainable, the paper details the best strategy used by the Sender. Similarly, Wei (2020) identifies how a Sender optimizes the signal for a Rational Inattention Agent. Interestingly, this paper shows that a little Rational Inattention is good for the Agent as it serves as a commitment that they will not pay attention if the Sender does not send more information than in the usual concavication case. Lipnowski et al. (2020) show where a benevolent Sender may choose how to restrict the information that a Rational Inattention Agent sees. Le Treust and Tomala (2019) have the Sender persuading the Agent by sending a message. However, in this case, the channel has limited capacity given by an upper bound to the reduction in the expected Shannon entropy. While similar, our approach differs from all these papers in two main ways: we remove the sending of a message (or signal) as a mean of persuasion, and we make endogenous the cost of obtaining information.

Strategic Obfuscation - A set of papers studying obfuscation motivated the economic problem that we detail here. When put together, the theoretical and empirical outcomes of this literature show that shrouding the truth can be a valid strategy. On a more theoretical note, Gabaix and Laibson (2006) develop a model in which companies can benefit from hiding the costs of add-ons (such as ink when selling printers) from myopic customers; Ellison and Ellinson (2009) analyze the use of obfuscation by Internet retailers as a way of dealing with how search engines reduce the cost of information. They test different mechanisms of obfuscation and detail the effects it has on the market. Ellison and Wolitzky (2012) give a theoretical model in which oligopolistic firms engage in price obfuscation to alter the equilibrium price of a homogeneous good. Richards et al. (2019) provide empirical evidence of obfuscation by firms trying to hinder price comparison on consumer goods by selling products with different sizes or attributes than their competitors. Kalayci and Potters (2011) and Jin et al. (2018) obtain outcomes of agents actively engaging in obfuscation, but in a laboratory environment. Ben-Shahar and Schneider (2014) extensively describe the possible problems with disclosure policies. More recently, Clippel and Rozen (2020) create and empirically test a model of communication in which obfuscation may appear in settings with mandatory disclosure. Obfuscation intensity is endogenous. They add strategic inference from the Agent, i.e., the Agent sees obfuscating behavior as a signal and draws inference from that.

3 The Model

There are two individuals, the Principal (*he*) and the Agent (*she*). The Agent must choose an action a from the set $A = \{a_H, a_L\}$. The set of payoff relevant states of nature is $\Omega = \{\omega_H, \omega_L\}$ and the realization of a state $\omega \in \Omega$ is drawn from a distribution $\mu_0 \in int(\triangle(\Omega))$, the prior, known by both individuals.[7] The Agent and Principal have material utility functions $u(a, \omega)$ and $v(a, \omega)$ respectively that depend on the chosen action.

Let $a^*_{Ag}(\mu) = argmax_{a \in A}\mathbb{E}_\mu[u(a, \omega)]$ represent the action that the Agent will take given a belief $\mu \in \triangle(\Omega)$, and we assume that no action is ever optimal for both states. For ease of exposition, we consider without loss of generality that $a_H(a_L)$ as the optimum action for the Agent when the state is $\omega_H(\omega_L)$. We can then deduce the existence of a $\mu_{ind} \in int(\triangle(\Omega))$ that represents the belief that makes the Agent indifferent between both actions. We will denote by $\mu \in [0, 1]$ the belief that the state is ω_H. Further, we assume that at the belief of indifference the Agent takes the action preferred by the Principal at that belief (Kamenica and Gentzkow, 2011).

To improve her odds of making the best decision, the Agent purchases a signal structure π that consists of signal realization set A that serves as a recommended action,[8] and a family of distributions $\{\pi(\cdot|\omega)\}_{\omega \in \Omega}$ over A. To purchase a signal, the Agent incurs in a cost which has

[7] We denote by $\triangle(X)$ the set of all probability distributions on a set X.

[8] In a rational inattention environment, it is without loss of generality to consider the signal realization set to be A and treat the signal realization as a recommended action. For more information, see Matějka and McKay (2015).

a magnitude chosen by the Principal. Here, we turn to one of the most ubiquitous ways of modeling the cost of information, as detailed in Sims (2003). Based on the expected reduction of Shannon Entropy (Shannon, 1948), the cost of a signal structure π is given by:

$$c(\pi; \mu_0, k) := k(H(\mu_0) - \mathbb{E}_\tau[H(\mu_a)]) \tag{1}$$

Where μ_a represents the posterior for a given signal realization $a \in A$; $\tau \in \triangle(\triangle(\Omega))$ is the distribution of posteriors induced by π; $k \in \mathbb{R}_+$ is the cost parameter. We assume that the Principal **chooses the value of** k: it is the only parameter the Principal has control over and it is the only way he can persuade the Agent to take a certain action. When there is an increase in k, gathering new information by purchasing an informative signal becomes more expensive. Lastly, H is Shannon Entropy. The Shannon Entropy function $H : \triangle(\Omega) \to \mathbb{R}_+$ for a given distribution μ is given by:

$$H(\mu) = -\sum_{\omega \in \Omega} \mu(\omega) ln(\mu(\omega))$$

Where for the extreme points we assume that the function takes the value of 0 as $lim_{x \to 0^+} x \cdot lnx = 0$, which is also in line with the literature (Matyskova and Montes, 2021; Caplin and Dean, 2013)). The intuitive idea of Shannon Entropy is that it can be interpreted as a measure of uncertainty and so the cost function is charging for how much uncertainty the signal reduces considering how much is known at the prior. Shannon Entropy holds many desirable properties: it induces economically interesting behaviors (for instance, knowing the state of the world for sure is prohibitively costly), it's mathematically simple, and it allows us to compare our results with others in the rational inattention literature that use the same function.

Upon choosing a signal structure, there will be a signal realization a. Since the Agent knows the conditional probability of seeing the signal realization in each of the possible states of the world, she updates her beliefs to μ_a according to Bayes' rule. She then takes action

$a^*_{Ag}(\mu_a)$, and utilities $u(a^*_{Ag}(\mu_a), \omega) - c(\pi; \mu_0, k)$ and $v(a^*_{Ag}(\mu_a), \omega)$ are realized.

Thus, the choice of k has an impact on the probability of the Agent taking a certain action and the Principal will then choose the value of k that maximizes his expected payoff. There are two ways a signal might not be costly: if the Principal sets $k = 0$ or if the Agent chooses an uninformative signal, and as a result $\mu_a = \mu_0$ and there is no change in entropy, i.e., the uncertainty is not reduced.

Here, we highlight a fundamental assumption in our model: the Principal chooses k before knowing the true state of the world. The idea behind this assumption is that, if the Principal knew with certainty the state, the choice of k itself would serve as a signal for the Agent. For instance, if the choice of k is too high, the Agent would understand that the Principal is seeking to confuse, which would imply the state of the world is not favourable for the Principal. The Agent would then take that into consideration as well before taking any action. To abstract from such discussions, we adopt the above mentioned assumption. Ultimately, it simply implies that there are uncertainties for both the Agent and the Principal, as neither have all the information.[9] In Section 5, we illustrate how this assumption works in more concrete examples.

3.1 The Agent's Problem

The Agent's problem consists in choosing an information structure based on the cost fixed by the Principal and a resulting action following the realization of the chosen information structure. Effectively, this boils down to a single problem to the Agent: she must choose which signal structure maximizes her expected utility (considering the costs). This is because after the signal structure is chosen, when the signal is realized she will simply perform the recommended action. There are two opposing forces in play at once: Agent wants to know the real state of nature so that she can choose the best possible action; Agent wants to learn

[9] In a similar fashion, Li and Shi (2017) provides a model in which a seller can allow the buyer to obtain more information about the valuation of a good. Neither the seller or the buyer have previous perfect information about the valuation about this information.

as little as possible to incur in less informational cost. From the interaction of these two forces, and in a similar fashion as in Caplin and Dean (2013), we have the following problem for the Agent.

Definition 1. *The Agent's Rational Inattention Problem is given by:*

$$
\begin{aligned}
& max_{\tau \in \triangle(\triangle(\Omega))} \mathbb{E}_{\tau}\left[U(\mu_a)\right] - k(H(\mu_0) - \mathbb{E}_{\tau}[H(\mu_a)]) \qquad (2) \\
& s.t.\ \mathbb{E}_{\tau}[\mu_a] = \mu_0
\end{aligned}
$$

Where $U(\mu_a) := \mathbb{E}_{\mu_a}[u(a^*_{Ag}(\mu), \omega)]$, and we have the usual restriction of Bayes Plausibility. Here it's important to highlight that we always assume that $\mu_0 \neq \mu_{ind}$, as then the Principal's Problem would not necessarily have a solution (this is an event of measure zero, but we explain in more detail why this is the case in Lemma 1 in the Appendix A).

Now, let τ_k denote the solution to (2) for a given k parameter.[10] Since we must obey Bayes' Plausibility - the restriction that the expected value of the posteriors must equal the prior - the solution τ_k also identifies the posteriors the Agent will reach given all possible signal realizations.

To identify these possible posteriors and characterize the solution, we call upon insights from Caplin and Dean (2013) and Matyskova and Montes (2021). Equation 3 comes from the *Likelihood Ratio Inequalities for Unchosen Acts*,[11] as described in Caplin and Dean (2013) or similarly in Lemma 3 of Matyskova and Montes (2021), that defines the regions in which a certain action $a \in A$ is always chosen:

$$
I(a) = \left\{ \mu \in \triangle\Omega : \sum_{\omega \in \Omega} \mu(\omega) \left(\frac{e^{\frac{u(a',\omega)}{k}}}{e^{\frac{u(a,\omega)}{k}}} \right) \leqslant 1,\ \forall a' \neq a \right\} \qquad (3)
$$

Recall that since in our environment we are dealing with two actions and two states, we identify a belief $\mu \in \triangle(\Omega)$ by a number in the interval $[0, 1]$ that represents the probability

[10] The solution is unique for two states and two actions. For more details see Caplin and Dean (2013)

[11] To give an intuition from where this solution comes from, it can be understood as a result of a first order condition on the Agent's problem. Matějka and McKay (2015) provides an analogous derivation.

that the state is ω_H.[12]

Equation 3 identifies two threshold beliefs that divide the interval $[0,1]$ of beliefs in three parts. Using the notation in our model, we can rewrite Equation 3 and define, for $k>0$, the threshold beliefs that identify the regions in which an action is taken with probability 1:

$$\underline{\mu}(k) := \frac{1-exp(\frac{u(a_H,\omega_L)-u(a_L,\omega_L)}{k})}{exp(\frac{u(a_H,\omega_H)-u(a_L,\omega_H)}{k})-exp(\frac{u(a_H,\omega_L)-u(a_L,\omega_L)}{k})} \tag{4}$$

$$\overline{\mu}(k) := \frac{1-exp(\frac{u(a_L,\omega_L)-u(a_H,\omega_L)}{k})}{exp(\frac{u(a_L,\omega_H)-u(a_H,\omega_H)}{k})-exp(\frac{u(a_L,\omega_L)-u(a_H,\omega_L)}{k})} \tag{5}$$

Furthermore, we take the following to ensure continuity:[13]

$$\underline{\mu}(0) := \lim_{k\to 0^+} \underline{\mu}(k) = 0$$

$$\overline{\mu}(0) := \lim_{k\to 0^+} \overline{\mu}(k) = 1$$

And, by calculating the limit when $k\to\infty$, we get:

$$\lim_{k\to\infty} \underline{\mu}(k) = \lim_{k\to\infty} \overline{\mu}(k) = \frac{u(a_L,\omega_L)-u(a_H,\omega_L)}{u(a_H,\omega_H)-u(a_H,\omega_L)+u(a_L,\omega_L)-u(a_L,\omega_H)} = \mu_{ind} \tag{6}$$

So when $k=0$ the thresholds are in the extremities, and as k increases, they move toward μ_{ind} at the interior. The two functions ($\underline{\mu}$ / $\overline{\mu}$) are strictly increasing/decreasing respectively[14] and they are continuous. As a consequence of these properties, we also have that $\mu_{ind} \in \left[\underline{\mu}(k), \overline{\mu}(k)\right]$ for all $k \in [0,\infty)$.

These two new functions and beliefs characterize the solution for the Agent. To put it simply, they divide the $[0,1]$ interval in three, as illustrated in Figure 1 for an arbitrary

[12]If $\mu_0 < \mu_{ind}$ and no new information is acquired, the Agent will choose action a_L as it maximizes the expected payoff at that belief. Analogously, if $\mu_0 > \mu_{ind}$ and no new information is acquired, the Agent will choose action a_H.

[13]To understand why we get these limits, note that $u(a_i,\omega_j)-u(a_j,\omega_j)<0$ and $u(a_i,\omega_i)-u(a_j,\omega_i)>0$ for $i\neq j$ by assumption.

[14]For more details, see Matyskova and Montes (2021). It is also possible to verify that the first derivative is always positive/negative.

value of k. It then all depends in which of the three sections the prior lies. If $\mu_0 \in [0, \underline{\mu}(k))$, then the cost of obtaining more information outweighs the benefits and the Agent prefers to purchase a non-informative signal (which costs 0), and learns nothing new. This implies that the Agent will always choose action a_L. Analogously, if $\mu_0 \in (\overline{\mu}(k), 1]$, the Agent buys a non-informative signal and always chooses a_H. If, however, $\mu_0 \in [\underline{\mu}(k), \overline{\mu}(k)]$, then there is gain in acquiring more information. The Agent will purchase a signal structure such that: if the signal realization (recommended action) is a_L the posterior will be $\underline{\mu}(k)$ and she chooses a_L; if the realization is a_H the posterior will be $\overline{\mu}(k)$, and she chooses a_H. In this last case, both actions are chosen with a positive probability. A direct consequence of this process is that as k grows and new information becomes more expensive, the interval in which purchasing new information is beneficial diminishes.

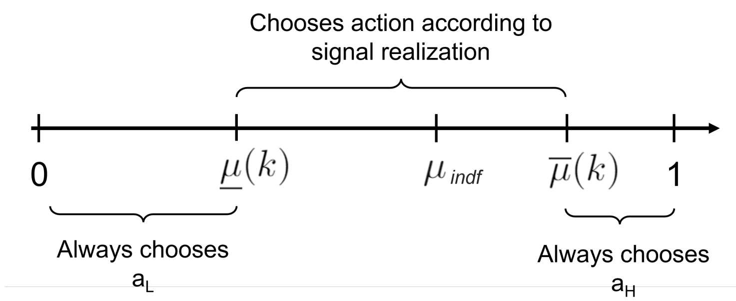

Figure 1: How the solution in Equation 3 divides the $[0, 1]$ interval. A belief indicates the probability that the state is ω_H.

To give a geometric intuition, these thresholds $\underline{\mu}(k)$ and $\overline{\mu}(k)$ identify the points where the indirect utility function minus the information cost detaches from its concavication. In Section 5 this is shown in more detail in the examples.

3.2 The Principal's Problem

Now that we know how the Agent will react, we are better equipped to understand how the Principal's utility changes with k. Again, the process is: Principal chooses k; Agent will choose an optimal distribution of posteriors τ_k (with support on $\underline{\mu}(k)$ and $\overline{\mu}(k)$ if information is purchased); this in turn implies an expected utility for the Principal.

To help us, we define the indirect expected utility given a belief for the Principal:

$$V(\mu) := \mathbb{E}_{\mu}[v(a^{*}_{Ag}(\mu), \omega)]$$

And, anticipating the optimal choice by the Agent, we can define the expected utility for the Principal given $k \in \mathbb{R}$:

$$\tilde{V}_{\mu_0}(k) := E_{\tau_k}[V(\mu_a)] \tag{7}$$

Where we use the prior as an index to highlight that the function changes with the prior. At this point, it is crucial to understand how this indirect utility works and, to this end, we will look at its geometric intuition. Take an illustrative case in which the prior is 75%, the payoff for the Principal is 0 if the Agent chooses a_L, 1 if the Agent chooses a_H, and $\mu_{ind} = 50\%$. With this, we can graph the function V for the Principal as in Figure 2.

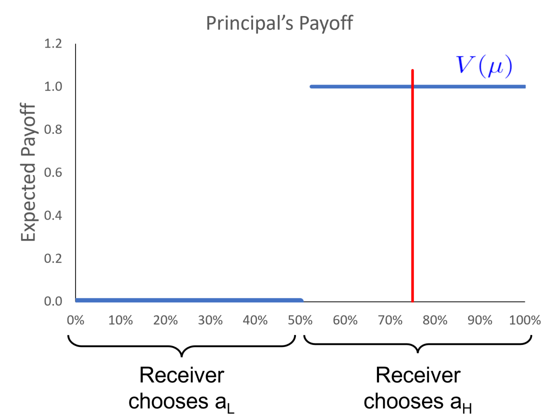

Figure 2: Expected indirect utility for the Principal given the belief. The horizontal axis denotes the probability that the state is ω_H.

With this, we are ready to try understand what happens when the Agent purchases a fully informative signal (when $k = 0$) and when the Agent purchases a non-informative signal (it happens when k is too high). Given that the Agent is subject to Bayes' Plausibility, - and this is very important - we can geometrically find the expected utility for the Principal at the intersection of the vertical line at the prior and the line connecting the indirect utility $V(\mu)$

evaluated at the posteriors.[15] This is more easily seen in Figure 3 on the left side, where we show the split of posteriors under full information. When a non-informative signal is acquired, the posterior will equal the prior and, similarly, we can find the expected utility for $k \to \infty$ by looking at the intersection between the vertical line at the prior and the function V, as shown on the right side of Figure 3.

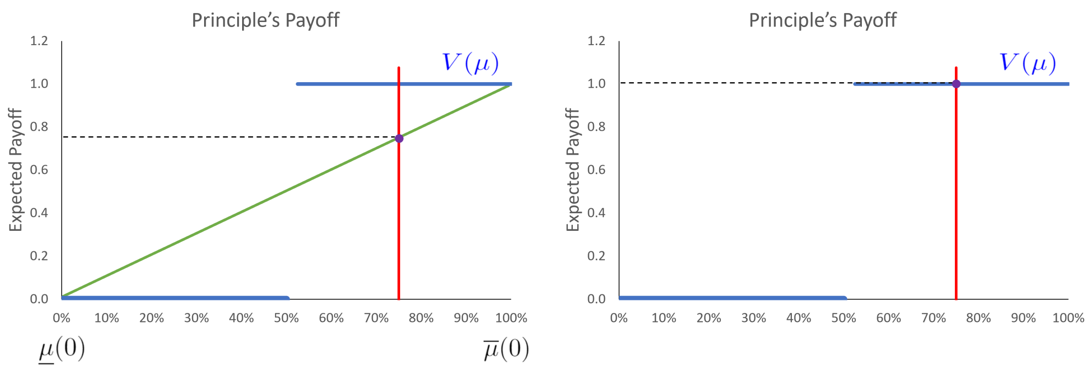

Figure 3: Left: the expected payoff of full information is at the intersection of the steep line with the vertical line of the prior. Right: the posterior is equal to the prior as k is too high. The Agent purchases a non-informative signal.

To illustrate the dynamics of how $\tilde{V}_{\mu_0}$ changes with k, consider Figure 4. As we've already seen, an increase in k moves the beliefs $\underline{\mu}(k)$ and $\overline{\mu}(k)$ closer to μ_{ind}. The value of $\tilde{V}_{\mu_0}$ then changes as the intersection of the vertical prior line with the line segment connecting V evaluated at the posteriors goes up or down. In the specific situation illustrated in the figure, the value of $\tilde{V}_{\mu_0}$ is increasing as the point of intersection is going up, but that is not always the case; this behavior depends on a number of factors including the preferences and the prior.

We can now deduce that there exists a value $\hat{k}$ high enough that can be defined by:

$$\begin{cases} \underline{\mu}(\hat{k}) = \mu_0, \text{ if } \mu_0 < \mu_{ind} \\ \overline{\mu}(\hat{k}) = \mu_0, \text{ if } \mu_0 > \mu_{ind} \end{cases} \tag{8}$$

To see this, first assume that $\mu_0 > \mu_{ind}$. Then, $\lim_{k\to\infty} \overline{\mu}(k) = \mu_{ind} < \mu_0 < 1 = \overline{\mu}(0)$.

[15]This geometric intuition is analogous to the concavication results in Kamenica and Gentzkow (2011).

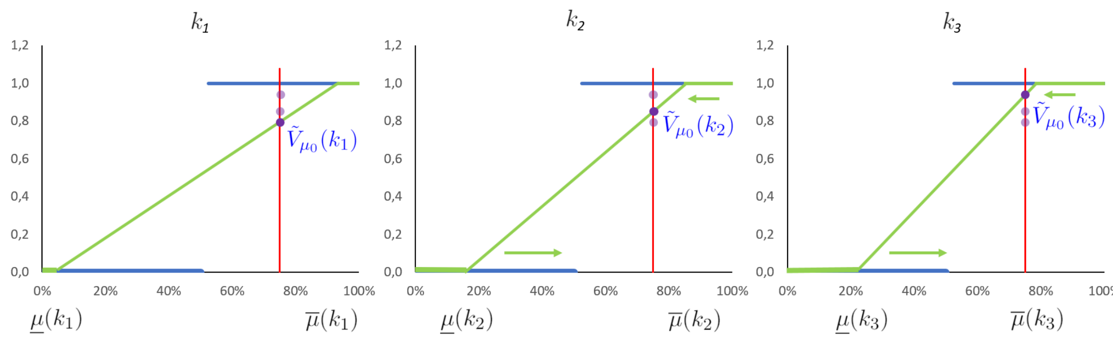

Figure 4: Dynamics of an increase in k. In this example, $k_1 < k_2 < k_3$.

Since $\overline{\mu}$ is strictly decreasing and continuous, as we increase k, by the Intermediate Value Theorem, there must be a unique value such that $\overline{\mu}(k) = \mu_0$. The logic is analogous if instead $\mu_0 < \mu_{ind}$. Economically, for $k \geqslant \hat{k}$, the Agent will only acquire non-informative signals and, as a result, will always choose action $a^*_{Ag}(\mu_0)$, as seen before. We then get that for all $k \geqslant \hat{k}$, $\tilde{V}_{\mu_0}(k) = V(\mu_0)$. Figure 5 illustrates a geometric interpretation for this fact.

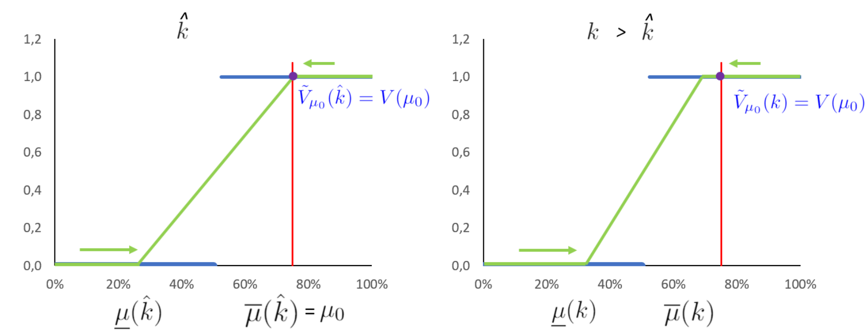

Figure 5: For $k > \hat{k}$, even if we increase k the intersection does not change, and $\tilde{V}_{\mu_0}(k)$ remains constant.

With this in mind, we find useful to define $B := [0, \hat{k}]$, the interval in which the Agent still acquires new information[16]. As we assume that $\mu_0 \in (0, 1)$, $\hat{k} > 0$ and B is not a degenerate interval. We can now describe a functional form for $\tilde{V}_{\mu_0}$. First, let us define an auxiliary function to help us with the notation:

$$m(k) := \left(\frac{V(\underline{\mu}(k)) - V(\overline{\mu}(k))}{\underline{\mu}(k) - \overline{\mu}(k)} \right)$$

[16] At $\hat{k}$ the Agent is in fact indifferent between purchasing information or not.

Then, the functional form is given by:

$$\tilde{V}_{\mu_0}(k) = \begin{cases} m(k)(\mu_0 - \underline{\mu}(k)) + V(\underline{\mu}(k)), \text{ if } k \in B \\ V(\mu_0), \text{ otherwise} \end{cases} \tag{9}$$

In other words, if k is too high ($k > \hat{k}$), the Agent plays according to the prior and $\tilde{V}_{\mu_0}(k) = V(\mu_0)$. If not, the value of $\tilde{V}_{\mu_0}(k)$ is given by the intersection of the line segment connecting the value of V at the posteriors in the support of τ_k evaluated at the prior.

Intuitively, by seeing the geometric interpretation, we can see that $\tilde{V}_{\mu_0}$ must be continuous as there can not be any jumps in the intersection as you increase k. But for a formal proof, it suffices to check the following cases. If $k \notin B$, then the function is constant; if $k \in B$ it is a composite function of linear combinations and divisions of continuous functions and it is therefore continuous.[17] We then need only to check for $k = \hat{k}$ as it is at this point that $\underline{\mu}(\hat{k}) = \mu_0$ or $\overline{\mu}(\hat{k}) = \mu_0$. If $\underline{\mu}(\hat{k}) = \mu_0$, simple substitution in Equation 9 yields $\tilde{V}_{\mu_0}(\hat{k}) = V(\underline{\mu}(\hat{k})) = V(\mu_0)$, and $\tilde{V}_{\mu_0}(k)$ is continuous at this point. The idea is analogous if $\overline{\mu}(\hat{k}) = \mu_0$.

Now that we have a better grasp on the workings of $\tilde{V}_{\mu_0}$, we have our Principal's Problem.

Definition 2. *The Principal's Obfuscation Problem is:*

$$\max_{k \in B} \tilde{V}_{\mu_0}(k) := \mathbb{E}_{\tau_k}[V(\mu_a)] \tag{10}$$

Note that if $\hat{k}$ is a solution to the Principal's Problem, then any $k > \hat{k}$ is also a solution, and so here we simplified things by assuming that in this case the Principal chooses $k = \hat{k}$. Finding $\hat{k}$, and therefore B, is just a matter of discovering k that satisfies:

$$m(k) * (\mu_0 - \underline{\mu}(k)) + V(\underline{\mu}(k)) = V(\mu_0)$$

[17] V may appear to not be continuous at μ_{ind}, but to reach this value we must be at $k > \hat{k}$, and at this value we will necessarily have $\tilde{V}_{\mu_0}(k) = V(\mu_0)$, which is constant and therefore continuous.

We discuss the solution to the Principal's problem in more depth in Sections 4 and 5.

4 Classifying The Principal's Problem

In this section we delve into the Principal's problem and detail the results we have found. In many situations, the optimal level of obfuscation is neither full disclosure nor full obfuscation, but rather, partial obfuscation.[18] We identify when partial obfuscation appears and discuss its possible implications.

Definition 3. *A Partial Obfuscation Principal's Problem is one in which neither* $k = 0$ *nor* $k = \hat{k}$ *are solutions for Equation 10. Otherwise it is called a Boundary Obfuscation Problem.*

We begin by looking at the shape of the Principal's indirect utility V. The proof of the next result can be found in Appendix A.

Proposition 1. *Given a Principal's Problem, we have:*

(i) If V *is concave, then* $\tilde{V}_{\mu_0}$ *is increasing in* k*.*

(ii) If V *is convex, then* $\tilde{V}_{\mu_0}$ *is decreasing in* k*.*

Proposition 1 serves as a tool for easily identifying that we always have a boundary problem if V is concave or convex. The indirect utility also provides a necessary condition for the problem to be of partial obfuscation by building on the fact that $\tilde{V}_{\mu_0}$ is twice differentiable in $int(B)$. Since this is the case, the first and second order condition apply, and we have that the problem is of partial obfuscation only if $\exists k^* \in int(B)$ such that $\frac{\partial \tilde{V}_{\mu_0}(k^*)}{\partial k} = 0$ and $\frac{\partial^2 \tilde{V}_{\mu_0}}{\partial k^2}(k^*) < 0$.

We now try to see how we can use the alignment between the Principal and the Agent as an effective way of checking if the problem has a boundary solution or not. It is economically

[18] Examples abound, such as the costs of add-ons, which are not completely hidden, but may be in fine print; or a firm that allows consumers to compare different models of mobile phones, but includes dozens of features to be checked, and so on.

intuitive that a higher cost is prejudicial when you want the Agent to know the state of the world and beneficial when you want the Agent to err. Let $a^*_P(\mu) = argmax_{a\in A}\mathbb{E}_\mu[v(a,\omega)]$ denote the optimal action for the Principal given a belief μ. We then describe the following cases:

Aligned Principal - Assume that $a^*_P(\mu) = a^*_{Ag}(\mu),\ \forall\mu \in \triangle(\Omega)$. In this case, the Principal's utility is decreasing in k for any prior as any mistakes on the part of the Agent are also sub-optimal for the Principal. The problem thus has a boundary solution. As a contrapositive from the Aligned Principal case, it is a necessary condition for the problem not to have a boundary solution at any given prior that there exists at least one belief μ in which $a^*_P(\mu) \neq a^*_{Ag}(\mu)$.

Misaligned Principal - Since we are in a two states / two actions environment, if $a^*_P(\mu) \neq a^*_{Ag}(\mu),\ \forall\mu \in \triangle(\Omega)$, we can already say that the Principal is better off by increasing k: whenever the Agent makes a mistake, she ends up choosing the optimal action for the Principal, and the problem has a boundary solution. For more complex environments with more actions, we would need a stronger assumption: $U(\mu) + V(\mu) = c,\ \forall\mu \in \triangle(\Omega)$, where c is a constant, effectively creating a zero sum game. Since the Agent cannot be better of by an increase in k, we directly get that the Principal's utility is increasing in k and we have a boundary obfuscation problem.

The next subsection focus on a particular kind of preferences that allows us to arrive at a full characterization of Partial Obfuscation Problems.

4.1 State Independent Preferences

We now turn our attention to the appealing class of problems in which the Principal has state independent preferences:[19] consider, for instance, a seller that does not care if their product is good or bad, only if the customer chooses to buy it or not. This is in fact a general case that has been studied in the literature of communication games (see, for example, Lipnowski

[19] We assume that one action is strictly preferred by the Principal at all states.

and Ravid (2020); Chakraborty and Harbaugh (2010)). Under this class of problems, we find necessary and sufficient conditions for having partial obfuscation. But first we begin with a simpler case.

Proposition 2. *If the Principal has State Independent Preferences and the Agent has Symmetric Preferences (i.e., $\mu_{ind} = \frac{1}{2}$), then we have a Boundary Obfuscation Problem.*

In such scenarios, the Principal chooses a boundary solution for any prior. Appendix A gives a detailed proof, but the main idea is that, in this environment, any increase in k will always result in symmetric movements by $\underline{\mu}$ and $\overline{\mu}$ and V is a step function, so that $\tilde{V}_{\mu_0}$ is monotonic.

From this result, we get the motivation for looking into the Agent from the perspective of the belief that makes her indifferent between the two actions. Indeed, μ_{ind} is uniquely defined by the Agent's utility function, and so it can be understood as an Agent's characteristic.

Definition 4. *Gullible Agent - We say that we have a Gullible Agent if:*

(i) The Principal's payoff is higher under action a_H, the indifference belief is lower than $\frac{1}{2}$ and the initial prior is even lower, i.e., $v(a_H, \cdot) > v(a_L, \cdot)$, with $\mu_0 < \mu_{ind} < 1/2$; or

(ii) The Principal's payoff is higher under action a_L, the indifference belief is higher than $\frac{1}{2}$ and the initial prior is even higher, i.e., $v(a_H, \cdot) < v(a_L, \cdot)$, with $\mu_0 > \mu_{ind} > 1/2$

The above definition helps us get to the main result for this section.

Theorem 1. *With State Independent preferences for the Principal, we have a Partial Obfuscation Problem if, and only if, we have a Gullible Receiver.*

The proof for the theorem can be found in Appendix A. We first prove that Gullible Receiver indeed yields a Partial Obfuscation problem, and then we proceed to show that all other possible cases result in Boundary Obfuscation problems.

The Gullible Receiver case arises when two situations occur at the same time: the Agent is in a situation in which choosing the action the Principal wants yields a high benefit in the right state, but not too much damage in the wrong state, so that obfuscation will induce the Agent to commit more mistakes; if choosing according to the prior, the Agent will take the least favorite action for the Principal. As an illustration, the Better or Worse Product example on section 5 falls under this case.

4.2 Comparison to Communication Games

We can compare our model to two canonical models of communication: Bayesian Persuasion (Kamenica and Gentzkow, 2011) and Cheap Talk (Crawford and Sobel, 1982). We keep the same environment of binary actions and binary states of the world.

The difference between the three models, namely: Strategic Obfuscation, Bayesian Persuasion and Cheap Talk can be framed in terms of how the Principal (Sender) can induce a posterior belief in the decision maker. In the Bayesian Persuasion environment, the Principal is able to commit to a signal structure while under Cheap Talk, the Sender is not. Under Strategic Obfuscation, the Principal decides how costly it is for the Agent to acquire new information.

Definition 5. *For a Communication Game with binary actions and binary states of the world, let s^P be the optimal payoff for the Sender under Bayesian Persuasion; let s^C be an equilibrium payoff for the Sender under Cheap Talk; and let s^O be the optimum payoff for the Principal under Strategic Obfuscation.*

With these in hand we have the following result.

Proposition 3. *Given a Communication Game, we have:*

(i) $s^O \leqslant s^P$*. Moreover, if we have an Aligned Principal,* $s^O = s^P$*.*

(ii) Under State Independent Preferences for the Principal, $s^C \leqslant s^O$*.*

The highest outcome that the Sender gets under Bayesian Persuasion is an upper bound to what can be achieved by the Principal under Strategic Obfuscation. On the other hand, for the class of state independent preferences, any equilibrium payoff for the Sender under Cheap Talk is weakly lower than the Principal's optimal payoff under Strategic Obfuscation. The proof is in Appendix A.

An intuition for this result is that in our model, the Principal can only control the intensity of the cost of information, so we are essentially restricting the ability of the Principal to induce signal structures. The only signals the Principal is able to induce for the Agent are the ones whose posteriors end up as $\underline{\mu}(k)$ and $\overline{\mu}(k)$ for some k, as defined in Equations 4 and 5. In Bayesian Persuasion, however, the Sender is free to choose any signal structure as long as it satisfies Bayes plausibility. Under Cheap Talk, the analysis is less straightforward and the result in Proposition 3 relies on the binary model and is restricted to the class of state independent preferences.

Our model is related to the one in Lipnowski et al. (2020): if we restrict our attention to consider a paternalistic Principal (in which $v = u$, for instance) we get that the best choice is full disclosure, or setting $k = 0$. In one of their results, Lipnowski et al. (2020) show that full disclosure is optimal if, and only if, there are two possible states of nature. Our model, however, could yield full disclosure as optimal even in environments with more states of nature. This happens because of a fundamental difference in the model: while in Lipnowski et al. (2020) they study how a Principal can choose the best information policy for a rationally inattentive agent to help them focus on the correct issues, our Principal can effectively always set the agent's cost of processing information to zero, so no mistakes happen even with more states.

5 Practical Examples

In this section we apply our model to practical economic contexts. Through this exercise, we provide interesting insights for agents and policy makers, specially regarding welfare. Note that many other economically relevant contexts - political, military, etc. - can be inserted in the model's framework.

5.1 Better or Worse Product

Producer, Inc. wants to sell their product to the Customer, who may choose between two actions: *buy* or *not buy*. The quality of the product can be $Worse$ or $Better$ than another similar product on the market (made by a different firm) that the Customer is already familiarized with. The common prior that the product is $Worse$ is given by $\mu_0 \in (0, 1)$. In accordance with our assumption, the prior is interior for both players because the Customer is not acquainted with all the features of the product and the Producer does not have all information on the Customer's preferences. The case of stealth startups can be applied here, as they are offering a new service/product that they are unsure if it will satisfy the market's needs. One can also think of the case of a re-seller, like Amazon, who was not present in the production process and so is unsure of the product's quality.

The Customer may realize a private investigation to obtain more information on the quality of the product: cost-benefit analysis, price comparison, search for online reviews, etc. The Producer may facilitate or hinder the quest for information by: selling the product in different sizes or batches than competitors, putting the specifications in fine print or not including the price for add-ons.[20]

If the Customer chooses *not buy* we assume that the payoff is zero for both players. If the Customer chooses *buy*, the Customer's payoff will be based on how much better the Product is compared to the competitors'. If it is $Better$, the benefit is given by $b > 0$, and if it is $Worse$, payoff is given by $-c$, where $c > 0$. The Producer gets a benefit of $d > 0$ independently of the state of the world if the Customer chooses *buy*. The payoffs are presented in Table 2, where rows denote the state of the world, not actions.

For a more concrete illustration, we set $\mu_0 = 0.75$, $d = b = 2$ and $c = 1$, which gives us

[20] Kalayci and Potters (2011) describes the ordeal of buying a mobile phone. With more than 30 listed technical attributes (such as weight, memory size and battery capacity), comparing options becomes an arduous task.

Table 2: Payoffs table for the Better or Worse Product game.

		Customer	
		not buy	buy
Producer	Better	$0, 0$	d, b
	Worse	$0, 0$	d, $-c$

that $\mu_{ind} \approx 0.667$. We then get the payoffs in Table 3. Note that we have a certain asymmetry considering the payoffs for when the Customer chooses *buy*. This can be interpreted as either the Customer not getting very angry when buying a worse product, or as there being in place a recycling policy that returns some benefit for the Customer if they choose to recycle the bad product.

Table 3: Payoffs table with numerical values.

		Customer	
		not buy	buy
Producer	Better	$0, 0$	2, 2
	Worse	$0, 0$	$2, -1$

Now, to solve the problem, we first plot U, the indirect utility for the Customer as a function of the beliefs, as seen on Figure 6. The horizontal axis denotes the probability that the state is *Worse*. On the right of Figure 6, it is shown the split of posteriors when $k = 0$. At this value of k the Customer will acquire a fully informative signal, and thus the posterior will be either 1 or 0. If, on the other hand, the cost is too high, the Customer will purchase a non informative signal, and will act according to the prior.

Since the beliefs are restricted by Bayes' Plausibility, we can obtain the expected utilities looking at where the vertical line of the prior intersects with the posteriors induced by the signal. When $k = 0$, we look at the intersection of the line that connect the points of the function U valued at the extremes 0 and 1 and the vertical line at the prior. When k is "too high" we look at the intersection of the vertical line at the prior (which equals the posterior in this case) and the function U. As a result, under full information the expected payoff is 0.5 and under no new information the expected payoff is 0.

Now that we know how the Customer will act at each belief, we can perform a similar

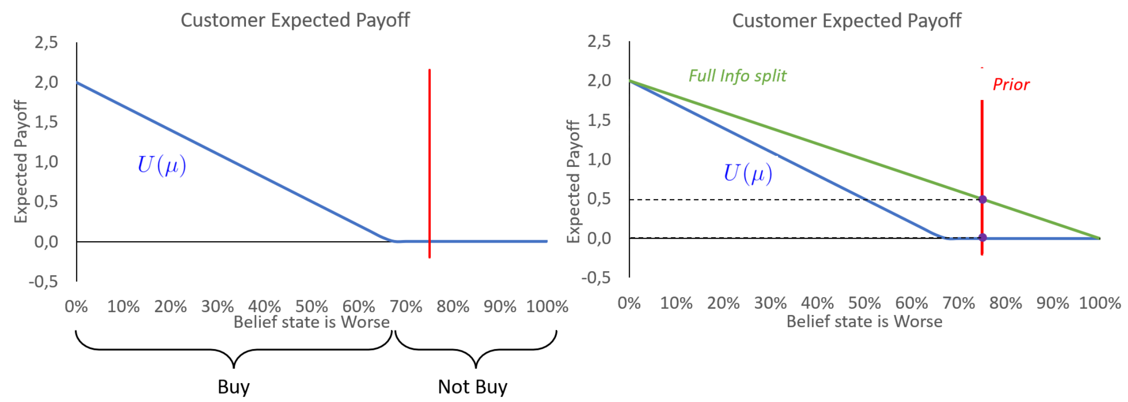

Figure 6: Payoffs as a function of beliefs for the Customer. Left: Curly brackets indicate the optimal action at that belief. Right: the diagonal green line indicates the split of posteriors and pay-offs for $k = 0$.

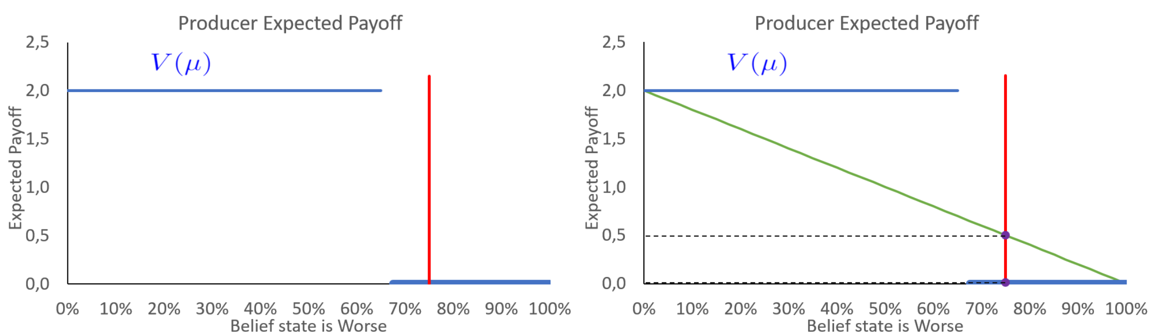

Figure 7: Payoffs as a function of k for the Customer. Right: the diagonal green line indicates the split of posteriors for $k = 0$.

analysis for the Producer. Figure 7 shows the expected utility V for the Producer. Again, on the right, we can find the expected payoffs for a given signal at the point where the vertical line of the prior intersects the induced split of posteriors. Under full information, the expected payoff for the Producer is 0.5; under no new information, the expected payoff is 0.

To look for the optimal Principal's obfuscation, we first illustrate how an intermediary k affects the Customer's choice of signal structure to purchase. Figure 8 plots the expected utility for the Customer on the left and the cost of information at each belief on the right. We chose for $k = 1$ to illustrate (keeping $\mu_0 = 0.75$). The sum of these two function results in the final function that the Customer will maximize (from 10).

We then can plot the function the Customer will maximize for $k = 1$ as shown in Figure

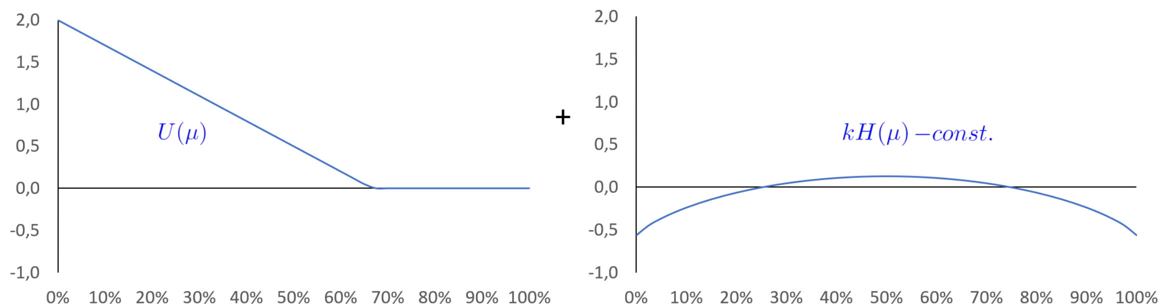

Figure 8: The function the Agent maximizes is given by the sum of these two functions. Left: indirect utility U. Right: Entropy for $k = 1$ and $\mu_0 = 0.75$.

9. The Consumer finds $\underline{\mu}(1) = 0.335$ and $\overline{\mu}(1) = 0.91$. Since the prior is between these two values, the Customer will learn by purchasing a signal that, when the realization is *buy*, updates his beliefs to $\underline{\mu}(1)$, and when the realization is *not buy*, updates his beliefs to $\overline{\mu}(1)$, so his expected payoff is 0.11. Had the prior fallen outside this interval, the Customer would have chosen a fully uninformative signal and acted according to the prior. Figure 9 also shows that $\underline{\mu}(k)$ and $\overline{\mu}(k)$ indicate where the concavication detaches from the expected utility U.

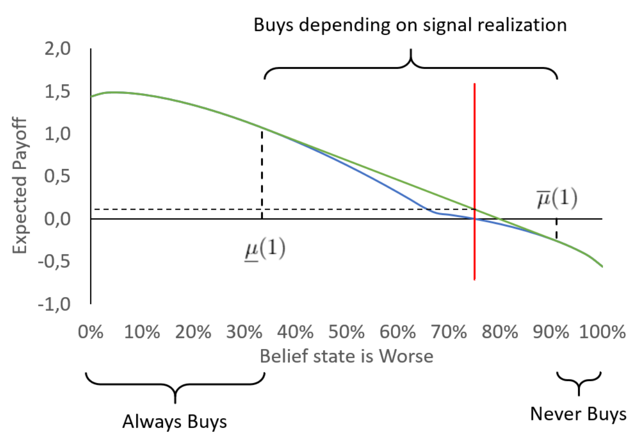

Figure 9: The maximizing function and its concavication for the Customer for $k = 1$ and $\mu_0 = 0.75$.

As an illustration of how an increase in k affects the maximizing function and the solution, Figure 10 shows on the left the maximizing function for $k = 0.5$, in which the solution yields a payoff of 0.243 and on the right the maximizing function for $k = 2$, which the solution yields a payoff of 0.02. It is visible that on the right the concavity of the Entropy is having a higher impact on the final function. We can also see that as k increases, $\underline{\mu}(k)$ and $\overline{\mu}(k)$ approach μ_{ind}.

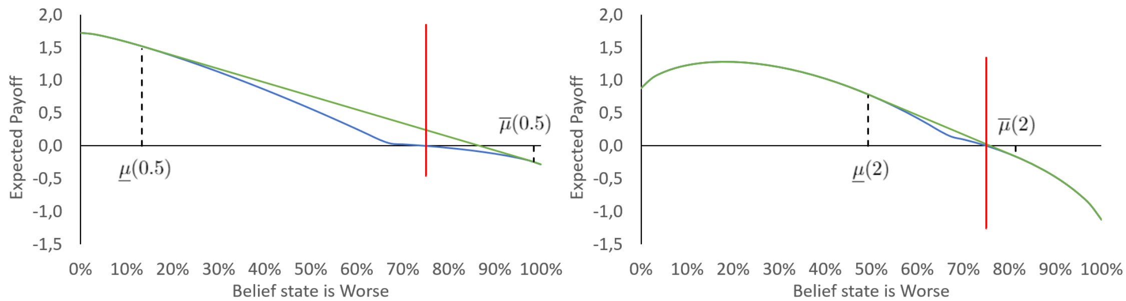

Figure 10: Customer's maximizing function and solution. Left: Case for $k = 0.5$. Right: Case for $k = 2$.

Now we return to the Producer and try to understand how intermediate values of k affect her payoffs. Figure 11 plots the expected payoff for the case in which $k = 1$. The payoff now is 0.556, higher than both the full information and no new information cases.

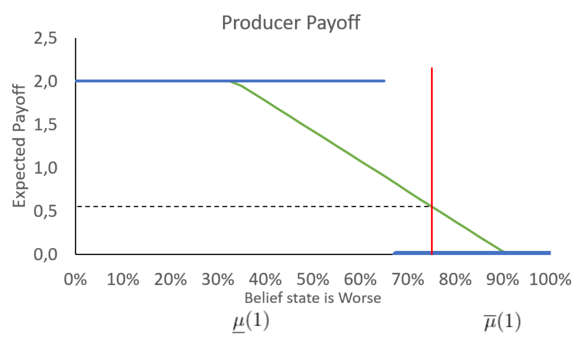

Figure 11: Producer's indirect utility and the chosen split for $k = 1$.

In fact, since we know the functional form for $\tilde{V}_{\mu_0}$, we can plot it for the different values of k. We do this in Figure 12. Our numerical solution shows that the optimal level of obfuscation ($k \approx 0.75$) indicates that the Producer should try to engage in just a little obfuscation. This means that the Producer may not divulge the price of all add-ons, but may clearly indicate the prices to facilitate comparisons with competitors. Just to note, in equilibrium we then have that the Customer gets an expected utility of 0.162 while the Producer gets an utility of 0.556.

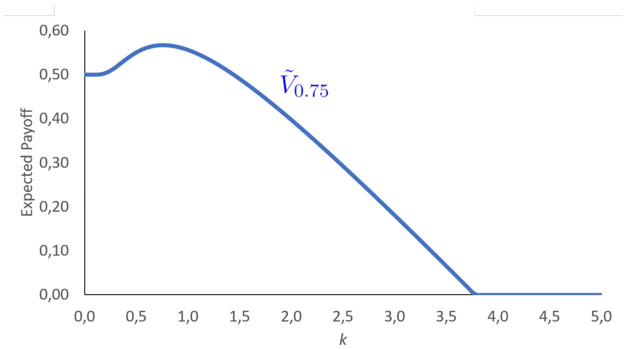

Figure 12: Producer's indirect utility from k.

Now, let's see what happens if we tweak the parameters a little bit. The only difference now is that we set $c = 2$. The idea is that we have more symmetry regarding the payoffs for the Customer when *buy* is chosen, which now makes $\mu_{ind} = 0.5$. The new payoffs table is shown in Table 4.

Table 4: Payoffs table with new value of c.

		Customer	
		not buy	buy
Producer	Better	$0, 0$	2, 2
	Worse	$0, 0$	$2, -2$

The steps to the solution in this new case are the same. We already know the behavior of the Customer as k increases, and we have the more general functional form of $\tilde{V}_{\mu_0}$ as given by Equation 9. Hence, we can plot the indirect utility for the Producer in this new context. Figure 13 shows the result. Now, the indirect utility is actually decreasing, and the optimal choice for the Producer is to engage in no obfuscation ($k = 0$), which has clear welfare implications.

Indeed, let k^* denote the equilibrium value of k, and we have the following expected payoffs in equilibrium: with $u(buy, Worse) = -1$, $k^* \approx 0.75$, the Producer has a payoff of 0.556 and the Customer has a payoff of 0.162; with $u(buy, Worse) = -2$, $k^* = 0$, the Producer has a payoff of 0.50 and the Customer has a payoff of 0.50.

When disregarding the obfuscating behavior, it seems that revoking the recycling policy

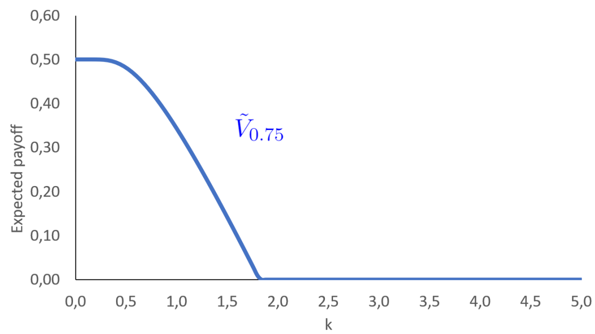

Figure 13: Producer's indirect utility from k for the new case in which $c = 2$.

would only make the Customer worse. We are effectively reducing the agent's utility in a given state with no apparent trade-off. However, as we've seen, this actually induces full transparency on the part of the Producer and, consequently, the Customer is actually better off in equilibrium. Another way of interpreting this is that if the Consumer could commit to "hurting" themselves if they are in the bad state, they could also induce the Principal to choose full disclosure. This fact highlights that policy makers and agents can greatly benefit from understanding the causes of obfuscating behavior in any economic problem.

In this example, the mechanism behind this phenomenon is the asymmetry in the Customer's payoffs as it affects the dynamics of the problem. It implies that the Customer's response to an increase in information costs will not be symmetrical when gathering information. When we set $c = 1$, we are making the Customer less averse to buying. As a result, the Customer is more inclined to purchase a signal structure that recommends *buy* more often, even if it leads to more mistakes. And, while k is still low, the Customer has enough leeway to choose this sort of signal. The Producer takes advantage of this fact and induces the Customer to *buy* more. As k increases however, this kind of signal becomes impracticable because, to recommend *buy* more and more frequently, the signal must also recommend *not buy* less and less, which requires accuracy, and that's expensive. At a certain point, the costs outweigh the benefits, and the Customer will be inclined to follow the prior, in which *not buy* is the better option. By setting $c = 2$, we are effectively making the Customer more "risk averse" and she will not be inclined anymore to choose a signal biased towards recommending *buy*: there is no space for exploiting the Customer.[21] The best course of action for the Producer is then full transparency to at least induce *buy* when the state is actually *Better*.

Finally, we also point out that depending on the parameters of the model, the original problem can be a partial obfuscation or a boundary obfuscation problem. For instance, if the prior was 0.25 (instead of 0.75), Figure 14 shows that now the best result is total obfuscation, or $k = \hat{k}$. What changed? The idea is that at the new prior of 0.25, the default action (i.e. if the Agent plays according to the prior) is *buy*. Thus, the Principal (the Producer) does not want the Agent to learn anything new and the Principal chooses maximum confusion.

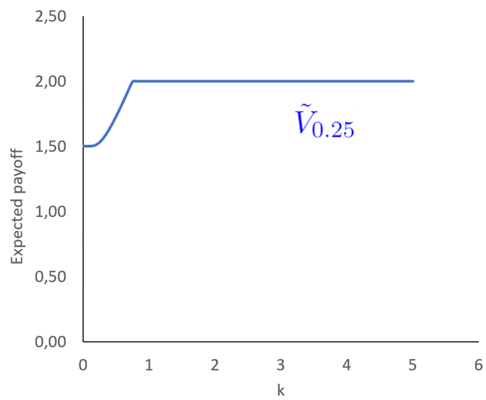

Figure 14: The indirect utility for the Principal in the Better or Worse Product example when the prior is 0.25.

5.2 Monopoly

Company Monopolistic Enterprises (MON - the incumbent) is a monopolist. Realizing the potential for profits, Competition Group (COM - the entrant), is deciding on whether they should make a play to enter this market. As a result, COM has to choose between two possible actions: *compete* and *not compete*. The success of COM's endeavor hinges on how strong MON's hold on the market is. We can interpret this as product quality, customer loyalty, capacity for further investment, etc. We denote by *Stronger* the state of the world

[21] This is an analogous result to the one in Martin (2017), albeit not in a pricing context, in which an increase in search costs simply induces the buyer to take the outside option, which is bad for the seller.

in which MON has a good hold on the market and COM's efforts will be unsuccessful; and by *Weaker* the state of the world in which MON does not have a good hold, and COM will be able to completely kick it off the market. The common prior that the state is *Weaker* is 0.7. The uncertainty for COM can be interpreted as not knowing for sure MON's strength. Similarly, MON is not sure how many resources COM can raise in its efforts to compete.

The payoff of being a monopolist in this market is given by $b > 0$, and the payoff of losing or not participating in this market is 0. Furthermore, we assume that COM has to pay a cost $d_C > 0$ to enter the market(for instance, investing in marketing, stores, hiring employees, etc.). To deal with the extra competition, MON incurs in a cost $0 < d_M < d_C$ if COM decides to enter.

COM would like to gather information on whether MON is indeed able to resist or not. Being rationally inattentive, they must bear a cost to do so. In our framework, MON does not send a signal to COM, i.e., it does not directly send information. Instead, it can set the parameter k that has an impact on how costly it is for COM to obtain further information. This can be interpreted as how closely they guard information (quarterly results, consumer reports, etc.) or if they release information (like rumours that they are about to open new factories or find new investments), etc.

The payoffs are explicitly shown in Table 5, where rows denote the possible states of the world instead of actions.

Table 5: Payoffs table for the Monopolist game.

		Competition Group	
		Not Compete	Compete
Monopolistic Enterprises	Stronger	$b, 0$	$b - d_M, -d_C$
	Weaker	$b, 0$	$-d_M, b - d_C$

For a more concrete example, in Table 6 we present the same payoffs table for the case in which we set $b = 1$, $d_C = 0.5$ and $d_M = 0.1$. Under these values, $\mu_{ind} = 0.5$.

Table 6: Payoffs table for the Monopolist game with values for the given parameters.

		Competition Group	
		Not Compete	Compete
Monopolistic Enterprises	Stronger	1, 0	0.9, -0.5
	Weaker	1, 0	-0.1, 0.5

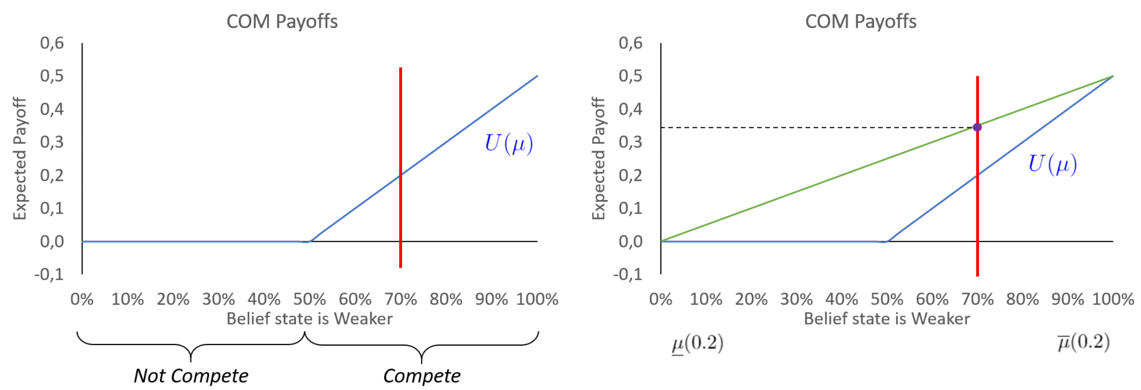

Figure 15: Payoffs as a function of beliefs for COM. The vertical red lines indicate the prior. Left: Curly brackets indicate the optimal action at that belief. Right: the diagonal green line indicates the split of posteriors for $k = 0$.

For the payoff structure in Table 6, Figure 15 shows the expected payoffs for COM as a function of the beliefs. The horizontal axis represent the belief that the state is *Weaker*. Since the agent chooses the action that maximizes the expected payoff at each belief, we were able to plot the indirect utility U for COM. On the right side of Figure 15 we also plot the split of posteriors under full information ($k = 0$) and it is straightforward that in this case the posterior will be either at 0 or at 1 (COM will try to know the state of the world with certainty). Given Bayes' Plausibility, we can find the payoff for full information at the intersection of the line connecting the value of U at the extremes and the vertical line at the prior. In the case of no new information, the posterior equals the prior and at the intersection of U with the vertical line at the prior we get the expected payoff with a non-informative signal. The payoff under full information is 0.23 and under no new information is 0.2.

Analogously, Figure 16 shows V for MON. Again, the diagonal line indicates the split of posteriors when COM has free access to a fully informative signal. Given the restriction of Bayes' Plausibility we get that he expected utility under full information is 0.23 and with

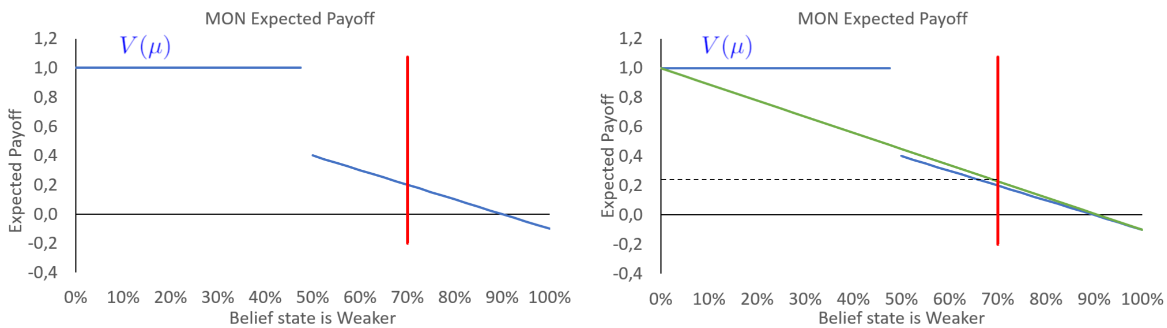

Figure 16: Payoffs as a function of beliefs for MON. The vertical red lines indicate the prior. On the right, the diagonal green line indicates the split of posteriors for $k = 0$.

no new information the expected payoff is 0.2. At first it seems that no obfuscation is the best option for MON, but it turns out that MON can do better. If the prior laid on the interval $(0, 1/2)$, the result would be rather simplistic. MON would like COM to get the least information possible, so that it would play according to the prior and choose *not compete*. However, since our prior is 0.7, we have an interesting situation. Indeed, Figure 17 might help us understand why. We first plot the function COM maximizes (as in Equation 10) when faced with a cost of $k = 0.2$. Again, we can see in the figure that the cost ends up "curving" the final function as a direct result of the concavity of the Entropy function.

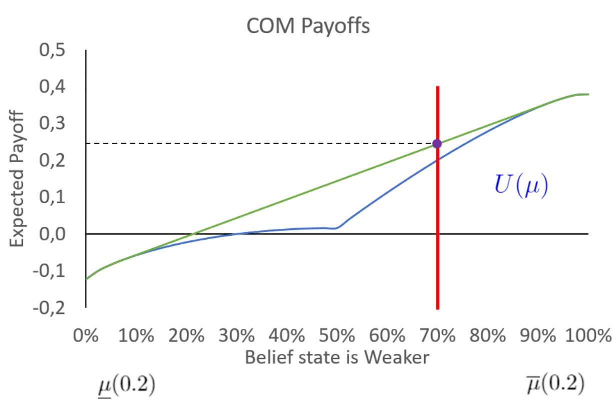

Figure 17: COM maximization function for $k = 0.2$. The split is at $\underline{\mu}(0.2) = 7.6\%$ and at $\overline{\mu}(0.2) = 92.4\%$.

As we've seen in Section 3.3.1, we first calculate $\underline{\mu}(0.2) = 7.6\%$ and $\overline{\mu}(0.2) = 92.4\%$. Since the prior is between these two values, COM buys an informative signal, i.e., there is gain in information, and the expected payoff for COM is 0.24. The interesting bit, however, appears when we graph the new split of posteriors at $k = 0.2$ for MON as in Figure 18. At this new split of posteriors, the expected payoff for MON is 0.25

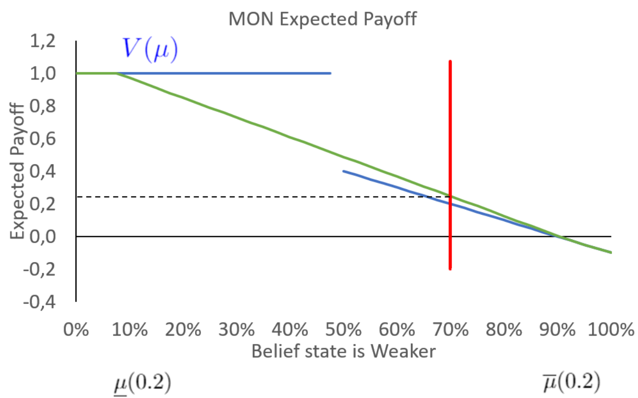

Figure 18: MON expected payoffs from the split of posteriors induced by $k = 0.2$. Expected utility is 0.25.

We then plot $\tilde{V}_{0.7}$. As plotted in Figure 19, our numerical solution for the optimal k is approximately 0.3. The reason why is that the payoffs for the monopolist when the competitor decides to compete depend on the state of the world. As such, by increasing the cost just a little, they may induce the competitor to enter the market when the state is *Stronger*, which is better than when they compete only when the state is *Weaker*. This also implies that the competitor will choose *not compete* when the state is *Weaker*, but that is not a problem for the monopolist, as in this action the payoff does not depend on the state. But if MON keeps increasing the value of k, CON will choose *compete* more and more, as the *compete* signal will appear with higher probability. After a certain point, costs will outweigh the benefits as the monopolist expects to be at the *Weaker* state most of the time (the prior is higher than 0.5).

And so, in equilibrium, MON will seek to confuse CON, but “just a little bit”. This is an example with State Dependent preferences in which we do have a Partial Obfuscation problem.

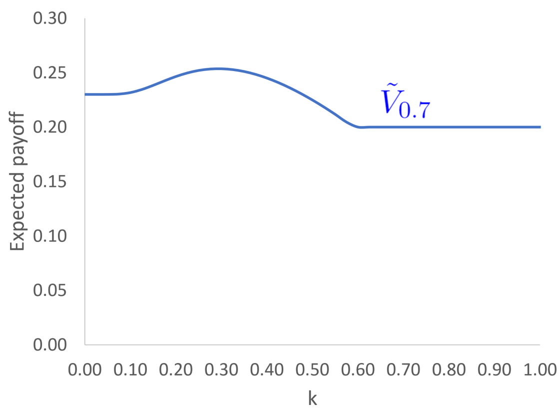

Figure 19: Plot of $\tilde{V}_{0.7}$.

6 Conclusions

In this paper we develop a model in which a Principal can induce particular actions on a rationally inattentive Agent by choosing the cost of processing information. We interpret that one can be persuasive by facilitating or hindering the access and processing of information.

First, we provide a new theoretical justification for why agents may rationally shroud information. This is in agreement with empirical papers on strategic obfuscation. In many cases, the optimal induced obfuscation level is neither zero nor "very high" but "just a little": the agent partakes in what we called *partial obfuscation*. We provide examples in which different mechanisms cause partial obfuscation to appear (for instance, an asymmetry in the Agent's preferences). There is a vast array of economic contexts in which the model can be applied and partial obfuscation problems may arise: competition, military, insurance, and politics. We have also characterized the partial obfuscation problems under the class of problems in which the Principal has state independent preferences.

Second, we found that obfuscation behavior has clear welfare implications as it can cause the Agent to err more. As seen in our Better or Worse Product example, if policy makers make decisions without fully comprehending the causes of obfuscation, they risk implementing ineffective or even harmful policies. Our model helps by shedding light into how some of the mechanisms that induce obfuscation can work.

Finally, a contribution of our paper is to provide a framework in which a Principal persuades an Agent only by controlling her access to information. Persuasion might be about information selection and information processing without having to restore to the strong commitment assumption in information transmission that is present in the canonical models of Bayesian Persuasion.

Given the nature of the model, it is open to many extensions, such as generalizing the results for more comprehensive state and action spaces and the implications of what happens when other cost structures are used. While we have used the most common cost function, recently the literature has started to study the implications of using other ways of modeling rational inattention. Caplin et al. (2022) categorize families of cost functions that may be applied in this context.

References

Alonso, R. and O. Camara (2022). Organizing data analytics. *Working paper*.

Barros, L. (2021). Information acquisition design. *Master Thesis, Sao Paulo School of Economics - FGV*.

Ben-Shahar, O. and C. E. Schneider (2014, apr). *More Than You Wanted to Know*. Princeton University Press.

Bergemann, D. and S. Morris (2019, March). Information design: A unified perspective. *Journal of Economic Literature 57*(1), 44–95.

Bizzotto, J., J. Rüdiger, and A. Vigier (2020). Testing, disclosure and approval. *Journal of Economic Theory 187*, 105002.

Bloedel, A. W. and I. Segal (2020). Persuading a Rationally Inattentive Agent. *Working Paper*.

Caplin, A. and M. Dean (2013). Behavioral Implications of Rational Inattention with Shannon Entropy. *NBER Working Paper* (August), 1–40.

Caplin, A. and M. Dean (2015). Revealed preference, rational inattention, and costly information acquisition. *American Economic Review 105*(7), 2183–2203.

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.

Chakraborty, A. and R. Harbaugh (2010). Persuasion by cheap talk. *American Economic Review 100*(5), 2361–82.

Clippel, G. D. and K. Rozen (2020). Communication , Perception and Strategic Obfuscation. *Working Paper*.

Crawford, V. P. and J. Sobel (1982, nov). Strategic Information Transmission. *Econometrica 50*(6), 1431.

de Oliveira, H., T. Denti, M. Mihm, and K. Ozbek (2017). Rationally inattentive preferences and hidden information costs. *Theoretical Economics 12*(2), 621–654.

Edmond, C. and Y. K. Lu (2017). Creating Confusion: Information Manipulation and Imperfect Coordination. (December).

Ellis, A. (2018). Foundations for optimal inattention. *Journal of Economic Theory 173*, 56–94.

Ellison, G. (2005, may). A Model of Add-On Pricing. *The Quarterly Journal of Economics 120*(2), 585–637.

Ellison, G. and S. F. Ellinson (2009). Search, Obfuscation, and Price Elasticities on the Internet. *Econometrica 77*(2), 427–452.

Ellison, G. and A. Wolitzky (2012). A search cost model of obfuscation. *RAND Journal of Economics 43*(3), 417–441.

Fréchette, G. R., A. Lizzeri, and J. Perego (2022). Rules and commitment in communication: An experimental analysis. *Econometrica 90*(5), 2283–2318.

Gabaix, X. and D. Laibson (2006, may). Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets. *The Quarterly Journal of Economics 121*(2), 505–540.

Gentzkow, M. and E. Kamenica (2014). Costly Persuasion. *American Economic Review 104*(5), 457–462.

Jin, G. Z., M. Luca, and D. Martin (2018). Complex Disclosure. *SSRN Electronic Journal*, 1–20.

Kalayci, K. and J. Potters (2011). Buyer confusion and market prices. *International Journal of Industrial Organization 29*(1), 14–22.

Kamenica, E. (2019). Bayesian persuasion and information design. *Annual Review of Economics 11*(1).

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

Le Treust, M. and T. Tomala (2019). Persuasion with limited communication capacity. *Journal of Economic Theory 184*, 104940.

Li, H. and X. Shi (2017, November). Discriminatory information disclosure. *American Economic Review 107*(11), 3363–85.

Lin, X. and C. Liu (2022). Credible persuasion. *Working paper*.

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

Lipnowski, E. and D. Ravid (2020). Cheap talk with transparent motives. *Econometrica 88*(4), 1631–1660.

Lipnowski, E., D. Ravid, and D. Shishkin (2022). Persuasion via weak institutions. *Journal of Political Economy 130*(10), 2705–2730.

Martin, D. (2017). Strategic pricing with rational inattention to quality. *Games and Economic Behavior 104*, 131–145.

Matějka, F. and A. McKay (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review 105*(1), 272–298.

Matyskova, L. and A. Montes (2021). Bayesian Persuasion with Costly Information Acquisition. *SSRN Electronic Journal* (178715).

Min, D. (2021). Bayesian persuasion under partial commitment. *Economic Theory 72*, 743–764.

Nguyen, A. and T. Y. Tan (2021). Bayesian persuasion with costly messages. *Journal of Economic Theory 193*, 105212.

Perez-Richet, E. and V. Skreta (2022). Test design under falsification. *Econometrica 90*(3), 1109–1142.

Piccione, M. and R. Spiegler (2012, 2). Price competition under limited comparability. *Quarterly Journal of Economics 127*, 97–135.

Richards, T. J., G. J. Klein, C. Bonnet, and Z. Bouamra-Mechemache (2019). *Strategic Obfuscation and Retail Pricing*. Number 0123456789. Springer US.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal 27*(4), 623–656.

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

Wei, D. (2020). Persuasion under costly learning. *Journal of Mathematical Economics*, 1–26.

A Appendix

Lemma 1 If $\mu_0 \neq \mu_{ind}$ then the Principal's Problem has a solution. If not, then there may not be a solution.

Proof. First we show that if $\mu_0 = \mu_{ind}$, there may be no solution. When $\mu = \mu_{ind}$ we must have that $\hat{k} = \infty$ by putting together Equation 6 and Equation 8. This in turn implies that $B = [0, \infty)$ which is not compact and thus we cannot invoke Weierstrass to say that a solution always exists for the Principal's Problem. We can go further and see that if we are in a case where the function $\tilde{V}_{\mu_0}$ is strictly increasing in B (for instance when we have a Misaligned Principal), the Principal would want to set $k = \infty$.[22] Figure 20 offers a geometric interpretation for when such a situation might arise.

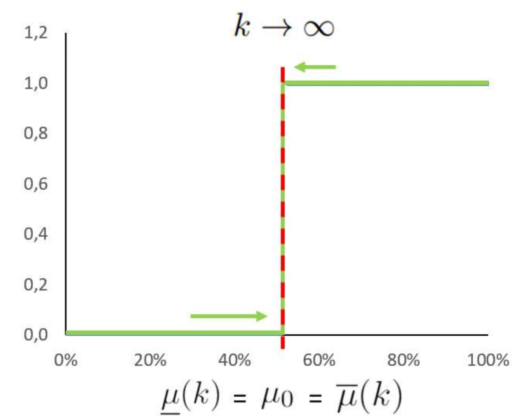

Figure 20: Illustration of may happen when $\mu_0 = \mu_{ind}$. The split of posteriors is shown for when $k \to \infty$.

For all priors $\mu_0 \neq \mu_{ind}$ however, $\hat{k} \in \mathbb{R}$ so that $B = [0, \hat{k}]$ is a compact set, and as $\tilde{V}_{\mu_0}$ is continuous, we can apply Weierstrass and the Principal's Problem has a solution. $\square$

Proof of Proposition 1

Given a Principal's Problem, we have:

(i) If V is concave, then $\tilde{V}_{\mu_0}$ is increasing in k.

[22] $\tilde{V}_{\mu_0}$ is not well defined as a function for $k = \infty$ as in that case $\underline{\mu}(k) = \overline{\mu}(k)$.

(ii) If V is convex, then $\tilde{V}_{\mu_0}$ is decreasing in k.

Proof. We know that there exists a $\hat{k}$ large enough such that, for every $k \geqslant \hat{k}$, $\tilde{V}_{\mu_0}(k) = V(\mu_0)$. As such, we will restrict our analysis for $k \in B = [0, \hat{k}]$. In this interval, $\tilde{V}_{\mu_0}(k) = m(k) * \mu_0 + V(\underline{\mu}(k)) - m(k) * \underline{\mu}(k)$. We shall detailed the proof only for the case in which V is concave as the proof for V convex is analogous. We then have that:

(i) $\tilde{V}_{\mu_0}$ valued at a specific $k \in [0, \hat{k}]$, is represented by the line segment connecting the points $(\underline{\mu}(k), V(\underline{\mu}(k)))$ and $(\overline{\mu}(k), V(\overline{\mu}(k)))$ valued at μ_0;

(ii) $\underline{\mu}$ and $\overline{\mu}$ are strictly increasing and decreasing respectively. This means that, for $0 \leqslant k_1 < k_2 \leqslant \hat{k}$, we have $0 \leqslant \underline{\mu}(k_1) < \underline{\mu}(k_2) \leqslant \mu_0 \leqslant \overline{\mu}(k_2) < \overline{\mu}(k_1) \leqslant 1$, with at most one equality for $\underline{\mu}(k_2) \leqslant \mu_0 \leqslant \overline{\mu}(k_2)$.

Thus, if we prove the general result that, for $a < c \leqslant \mu_0 \leqslant d < b$ (again, with at least one strict inequality), and a concave function $f : I \to \mathbb{R}$, where $[a, b] \subset I$, the intersection of the line segment connecting $(a, f(a))$ and $(b, f(b))$, when evaluated at μ_0, has an equal or lower value than the line segment connecting $(c, f(c))$ and $(d, f(d))$ when evaluated at μ_0, we are done. To put it simply, we are considering that $[a, b]$ represents $[\underline{\mu}(k_1), \overline{\mu}(k_1)]$ and $[c, d]$ represents $[\underline{\mu}(k_2), \overline{\mu}(k_2)]$ for arbitrary $k_1 < k_2$.

Figure 21 gives a visual representation. Points A and B represent the intersection of different line segments. What we wish to prove is that Point B is lies at least as high as Point A, as Point B is the intersection of the line $y = \mu_0$ with the innermost line segment (which would correspond to $\tilde{V}_{\mu_0}$ at a given cost k).

We can write the formulas for each line segment, for $x_1 \in [a, b],\ x_2 \in [c, d]$ as the following functions:

$$
\begin{aligned}
C1(x_1) &: \left(\frac{f(b) - f(a)}{b - a}\right) x_1 + f(a) - \left(\frac{f(b) - f(a)}{b - a}\right) a \\
C2(x_2) &: \left(\frac{f(d) - f(c)}{d - c}\right) x_2 + f(c) - \left(\frac{f(d) - f(c)}{d - c}\right) c
\end{aligned}
$$

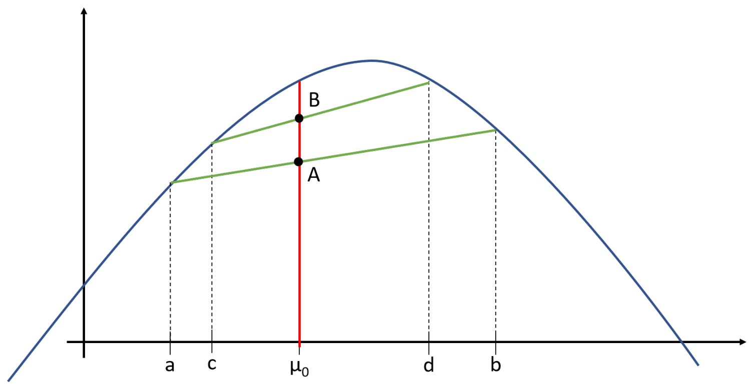

Figure 21: A concave function f and the intersection of the two line segments with the vertical line at μ_0.

Now, evaluating $C1$ and $C2$ at c, we get:

$$C1(c) : (f(b) - f(a))\,\frac{(c-a)}{(b-a)} + f(a)$$
$$C2(c) : f(c)$$

But since $c \in (a,b)$, there exists $t_c \in (0,1)$ such that $c = t_c b + (1-t_c)a$. We then get:

$$\frac{(c-a)}{(b-a)} = \frac{t_c b + a - t_c a - a}{b-a} = t_c \frac{b-a}{b-a} = t_c$$

And so:

$$(f(b) - f(a))\,\frac{(c-a)}{(b-a)} + f(a) = (f(b) - f(a))\,t_c + f(a) \tag{11}$$

But, since f is concave:

$$C2(c) = f(c) = f(t_c b + (1-t_c)a) \geqslant t_c f(b) + (1-t_c) f(a) = (f(b) - f(a))\,t_c + f(a) = C1(c)$$

The same argument can be done for the point d. Since $d \in (a,b)$, there exists $t_d \in (0,1)$ such that $d = t_d b + (1-t_d)a$, and we get that:

$$C2(d) = f(d) = f(t_d b + (1-t_d)a) \geqslant t_d f(b) + (1-t_d) f(a) = (f(b) - f(a))\,t_d + f(a) = C1(d)$$

And so, $C1(c) \leqslant C2(c)$ and $C1(d) \leqslant C2(d)$. Since the line segments are by definition linear,[23] $C2$ is never below $C1$ in the interval $[c, d]$. Finally, as $\mu_0 \in [c, d]$, then $C2(\mu_0) \geqslant C1(\mu_0)$. As mentioned before, $C2(\mu_0)$ represent the value of $\tilde{V}_{\mu_0}$ at k_2 and $C1(\mu_0)$ the value of $\tilde{V}_{\mu_0}$ at k_1, and we get that $\tilde{V}_{\mu_0}$ is increasing in k. $\square$

Proof of Proposition 2

If the Principal's preferences are state independent (i.e. they depend only on the action and not on the state) and the Agent's are symmetric (i.e. $\mu_{ind} = \frac{1}{2}$), then we have a Boundary Obfuscation Problem.

Proof. First, note that since the Agent's preferences are symmetric, $\underline{\mu}(k) = 1 - \overline{\mu}(k)$ for all k. This can be checked by noting that $\mu_{ind} = 0.5$ implies that $u(a_H, \omega_H) - u(a_L, \omega_H) = u(a_L, \omega_L) - u(a_H, \omega_L)$ and using substitution in Equations 4 and 5. Then, given that Principal's preferences are state independent, for any belief $\mu < \mu_{ind}$, $V(\mu) = V(0)$ and for any belief $\mu > \mu_{ind}$, $V(\mu) = V(1)$. Assume that $V(0) > V(1)$ (the other case follows a similar argument). Now, fix $k_1, k_2 \in B = [0, \hat{k}]$ (as outside this interval the function is constant) with $k_1 < k_2$, and let $d := \mu_0 - \underline{\mu}(k_1)$, $D := \overline{\mu}(k_1) - \mu_0$. These last variables represent the distance between the each of the induced posteriors and the prior.

By triangle similarity, we have:

$$\frac{d}{V(\underline{\mu}(k_1)) - \tilde{V}_{\mu_0}(k_1)} = \frac{D}{\tilde{V}_{\mu_0}(k_1) - V(\overline{\mu}(k_1))}$$

But since $\underline{\mu}(k_1) < \mu_{ind}$ and $\overline{\mu}(k_1) > \mu_{ind}$, we have $V(\underline{\mu}(k_1)) = V(0)$ and $V(\overline{\mu}(k_1)) = V(1)$. We can then write:

$$\frac{d}{V(0) - \tilde{V}_{\mu_0}(k_1)} = \frac{D}{\tilde{V}_{\mu_0}(k_1) - V(1)}$$

Simplifying:

$$\tilde{V}_{\mu_0}(k_1) = \frac{D}{d + D} V(0) + \frac{d}{d + D} V(1) \tag{12}$$

[23] This implies that either they are on the same line or they intersect at most once.

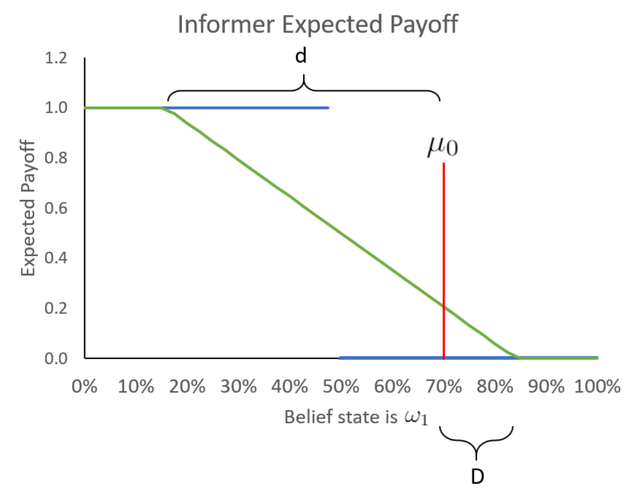

Figure 22: Simple visualization of the d and D variables for an illustrative State Independent and Symmetric case.

Now, let $\delta := \underline{\mu}(k_2) - \underline{\mu}(k_1)$. Since Agent's preferences are Symmetric, $\delta = \overline{\mu}(k_1) - \overline{\mu}(k_2)$, and we can write:

$$\tilde{V}_{\mu_0}(k_2) = \frac{D-\delta}{d+D-2\delta}V(0) + \frac{d-\delta}{d+D-2\delta}V(1) \tag{13}$$

But note that by construction, $\delta < \frac{d+D}{2}$, and so the following holds:

$$\frac{D-\delta}{d+D-2\delta} > \frac{D}{d+D} \iff d < D$$

We now have two cases:

(i) $D > d$: This happens when $\mu_0 < \mu_{ind}$. This also implies that from Equation 12 to Equation 13 we are increasing the proportion of $V(0)$ and since we assume that $V(0) > V(1)$, this means that $\tilde{V}_{\mu_0}(k_1) < \tilde{V}_{\mu_0}(k_2)$ and the problem has a boundary solution as $\tilde{V}_{\mu_0}$ is increasing.

(ii) $d > D$: This happens when $\mu_0 > \mu_{ind}$. This also implies that from Equation 12 to Equation 13 we are decreasing the proportion of $V(0)$ and since we assume that $V(0) > V(1)$, this means that $\tilde{V}_{\mu_0}(k_1) > \tilde{V}_{\mu_0}(k_2)$ and the problem has a boundary solution as $\tilde{V}_{\mu_0}$ is decreasing.

$\square$

Proof of Theorem 1

With state independent preferences for the Principal, we have a Partial Obfuscation Problem if, and only if, we have a Gullible Receiver.

Proof. Let the Principal have state independent preferences, with $v(a_H, \cdot) \neq v(a_L, \cdot)$. We call the Agent gullible if:

(i) $v(a_H, \cdot) > v(a_L, \cdot)$, with $\mu_0 < \mu_{ind} < 1/2$; or

(ii) $v(a_H, \cdot) < v(a_L, \cdot)$, with $\mu_0 > \mu_{ind} > 1/2$

We start by showing sufficiency, i.e., that with a Gullible Receiver the problem is always of partial obfuscation. Our goal is to find a k that yields a higher payoff for the Principal than either $k = 0$ or $k = \hat{k}$. We focus here on the case in which $v(a_L, \cdot) > v(a_H, \cdot)$, or equivalently, $V(0) > V(1)$, as the other case follows a symmetric argument. Then, if we have a gullible Agent, $\mu_{ind} > 1/2$ and $\mu_0 > \mu_{ind}$, which implies that the Agent would choose a_H if they act according to their prior. In other words, $\tilde{V}_{\mu_0}(\hat{k}) = V(1)$.

As usual, $B = [0, \hat{k}]$, and by triangle similarity, $\tilde{V}_{\mu_0}(0) = V(0)(1 - \mu_0) + V(1)\mu_0$. Given our assumption that $V(0) > V(1)$, for any $k \in int(B)$, $\tilde{V}_{\mu_0}(k) > \tilde{V}_{\mu_0}(\hat{k})$. So we must check if there exists a k in $int(B)$ such that $\tilde{V}_{\mu_0}(k) > \tilde{V}_{\mu_0}(0)$. Indeed, by triangle similarity, we have that for any $k \in int(B)$:

$$\tilde{V}_{\mu_0}(k) = V(0)\frac{\overline{\mu}(k) - \mu_0}{\overline{\mu}(k) - \underline{\mu}(k)} + V(1)\frac{\mu_0 - \underline{\mu}(k)}{\overline{\mu}(k) - \underline{\mu}(k)}$$

Thus, if we find a k such that

$$1 - \mu_0 < \frac{\overline{\mu}(k) - \mu_0}{\overline{\mu}(k) - \underline{\mu}(k)} \tag{14}$$

We are done. Now, note that both $\underline{\mu}$ and $(1 - \overline{\mu})$ are positive, continuous and strictly increasing functions with $\lim_{k \to 0^+} \underline{\mu}(k) = \lim_{k \to 0^+} 1 - \overline{\mu}(k) = 0$ and define $\lambda := \frac{\mu_0}{1-\mu_0}$, so that Equation 14 gives:

$$1 - \mu_0 < \frac{\overline{\mu}(k) - \mu_0}{\overline{\mu}(k) - \underline{\mu}(k)} \iff \underline{\mu}(k) > \lambda(1 - \overline{\mu}(k)) \iff 1 > \lambda \frac{(1 - \overline{\mu}(k))}{\underline{\mu}(k)} \tag{15}$$

Since $\mu_{ind} > 1/2$, we have that $M := u(a_L, \omega_L) - u(a_H, \omega_L) > m := u(a_H, \omega_H) - u(a_L, \omega_H)$, and this together with Equations 4 and 5 gives us that:

$$\lim_{k \to 0^+} \frac{1 - \overline{\mu}(k)}{\underline{\mu}(k)} = 0 \implies \lim_{k \to 0^+} \lambda \frac{1 - \overline{\mu}(k)}{\underline{\mu}(k)} = 0 \tag{16}$$

To help us understand the mechanics of this limit, note that using the functional form for $\underline{\mu}(k)$ and $\overline{\mu}(k)$ the limit becomes:

$$\lim_{k \to 0^+} \frac{1 - \overline{\mu}(k)}{\underline{\mu}(k)} = \lim_{k \to 0^+} \frac{1 - \frac{1-e^{\frac{M}{K}}}{e^{\frac{-m}{k}} - e^{\frac{M}{k}}}}{\frac{1-e^{\frac{-M}{K}}}{e^{\frac{m}{k}} - e^{\frac{-M}{k}}}} = \lim_{k \to 0^+} \frac{e^{\frac{-m}{k}} - 1}{e^{\frac{-m}{k}} - e^{\frac{M}{K}}} \frac{e^{\frac{m}{k}} - e^{\frac{-M}{k}}}{1 - e^{\frac{-M}{K}}} = 0$$

And the final equality is a direct consequence of $M > m$. Therefore, we deduce the existence of $k \in (0, \hat{k})$ small enough such that Equation 15 is satisfied, and we have a partial obfuscation problem.

We now prove necessity by checking all possible cases we can have with state independent preferences. Let $\mu_{ind} > \frac{1}{2}$. We then have the following possible cases:

1. $v(a_L, \cdot) > v(a_H, \cdot)$:

   (i) $\mu_0 > \mu_{ind}$: this is a Gullible Agent case and it is therefore a partial obfuscation problem.

   (ii) $\mu_0 < \mu_{ind}$: the optimum for the Principal is to choose $k = \hat{k}$ given that $a^*_{Ag}(\mu_0) = a_L$. The problem is thus of boundary obfuscation.

2. $v(a_L, \cdot) < v(a_H, \cdot)$:

   (i) $\mu_0 > \mu_{ind}$: the optimum for the Principal is to choose $k = \hat{k}$ given that $a^*_{Ag}(\mu_0) = a_H$. The problem is thus of boundary obfuscation. Note, however, that for analogous reasons to when we have a Gullible Agent, $\tilde{V}_{\mu_0}$ is non-monotonic in this case. The difference here is that the k that induces the lowest payoff is now interior. Indeed, confusing "just a little" might be the worst decision for the Principal.

   (ii) $\mu_0 < \mu_{ind}$: the optimum for the Principal is to choose $k = 0$ as in this case $\tilde{V}_{\mu_0}$ is decreasing in k, and so the problem is of boundary obfuscation. To see this, note that $\underline{\mu}(k) > 1 - \overline{\mu}(k)$. Building on the construction for the symmetric case proof and using Equation 12, we see that as k rises we are increasing the proportion of $V(0) = v(a_L, \cdot)$, which in this case we assume is the worst option.

The cases where $\mu_{ind} < \frac{1}{2}$ are analogous. Since we've already shown that when $\mu_{ind} = \frac{1}{2}$ the problem is of boundary obfuscation, the proof is complete. □

Proof of Proposition 3

Given a Communication Game, we have:

(i) $s^O \leqslant s^P$. Moreover, if we have an Aligned Principal, $s^O = s^P$.

(ii) Under State Independent Preferences for the Principal, $s^C \leqslant s^O$.

Proof. We start by proving (i). Note that in Bayesian Persuasion the Principal maximizes his utility by choosing a distribution of posteriors in the set:

$$BP(\mu_0) := \{\tau \in \triangle(\triangle(\Omega)) : \sum_{supp(\tau)} \mu\tau(\mu) = \mu_0\}$$

So, given commitment, the Sender is only restricted by Bayes plausibility. Under Strategic

Obfuscation, the Principal maximizes his utility in the set:

$$SO(\mu_0) \coloneqq \{\tau \in \triangle(\triangle(\Omega)) : \exists k \in B \text{ with } supp(\tau) \subseteq \{\overline{\mu}(k), \overline{\mu}(k)\} \wedge \sum_{supp(\tau)} \mu\tau(\mu) = \mu_0\}$$

We can clearly see that $SO(\mu_0) \subseteq BP(\mu_0)$, and so the Principal in Strategic Obfuscation cannot do better then the Sender in Bayesian Persuasion. When we have an Aligned Principal $a^*_P(\mu) = a^*_{Ag}(\mu),\ \forall \mu \in \triangle(\Omega)$ and under the Bayesian Persuasion framework, the best the Principal can do is to choose a signal structure that fully reveals the state. But that can always be achieved in Strategic Obfuscation by setting $k = 0$.

For the proof of (ii), we go back to the case of state independent preferences for the Principal. We repeat the claims in Proposition 2, and for any belief $\mu < \mu_{ind}$, $V(\mu) = V(0)$ and for any belief $\mu > \mu_{ind}$, $V(\mu) = V(1)$. Assume that $V(0) > V(1)$ (the other case follows a similar argument). From Lipnowski and Ravid (2020), we know that an outcome (τ, s) is an equilibrium outcome if, and only if, $\tau \in \triangle(\triangle(\Omega))$ satisfies Bayes plausibility and $s \in \cap_{\mu \in supp(\tau)} V(\mu)$. But, given the state independent preferences, we have that $\forall F \in 2^{\triangle(\Omega)}$, the following holds:

$$\begin{cases} \cap_{\mu \in F} V(\mu) & = \{V(0)\} \text{ if } F \subseteq [0, \mu_{ind}] \\ \cap_{\mu \in F} V(\mu) & = \{V(1)\} \text{ if } F \subseteq (\mu_{ind}, 1] \\ \cap_{\mu \in F} V(\mu) & = \emptyset \text{ otherwise} \end{cases}$$

And so the only possible payoffs for the Sender in a Cheap Talk equilibrium are $V(0)$ or $V(1)$. When adding Bayes Plausibility, we have one of two cases. If $\mu_0 < \mu_{ind}$, the only possible payoff is $V(0)$, and this can be achieved under Strategic Obfuscation by setting $k = \hat{k}$ (since $\mu_0 < \mu_{ind}$, the decision maker will take the optimum action for the Principal if no new information is acquired); if $\mu_0 > \mu_{ind}$, $V(1)$ is the only achievable payoff in equilibrium for the Sender, i.e., the Receiver chooses the worst action for the Sender with probability 1. But in Strategic Obfuscation the Principal can always get a better payoff by setting $k = 0$. Since μ_0 is interior, by fully revealing the state the optimal action for the Principal is expected to be chosen with strictly positive probability. $\square$