Swapping Signals: When and How Learning is Obfuscated

Srijita Ghosh[*]and Eric Spurlino[†]

February 2024

Abstract

In this paper, we consider a market where information is readily available but often cognitively costly and sellers can directly affect the learning,i.e., obfuscate information for the buyer in myriad of ways. Using the framework of a one-shot strategic communication game, we model the equilibrium obfuscation behavior of the seller. In our model, the buyer pays a cognitive cost of learning and the seller can garble the posterior belief distribution of the buyer directly. We find that in equilibrium, if the buyer's belief is favorable it is optimal to obfuscate fully, however, in case of unfavorable belief an intermediate (or zero) level of obfuscation becomes optimal. We also find that the range of parameters where obfuscation is optimal expands with the cognitive capacity of the buyer. Our framework is agnostic about the form of obfuscation. Furthermore, we use two examples, namely, hiding information and providing misleading information to demonstrate how our model can be used to make testable predictions across different obfuscation practices.

1 Introduction

Most consumer purchase decisions come after some degree of information acquisition—whether it's comparing multiple options for a desired good, introspecting on one's preferences, or understanding the total cost and affordability of the good for sale. The nature of commerce in the Internet age means that (1) this information is readily available but often cognitively costly to learn, and (2) sellers can directly affect this learning, i.e., *obfuscate* the buyer's learning efforts in a myriad of ways. Dark patterns, misleading advertising, and direct lying (or fraud) are just some of the ways online sellers affect the learning abilities of their consumers.

The economic literature on obfuscation can be broadly divided into three styles. In one, sellers are given the ability to manipulate the search costs of the buyer, increasing the time it takes to search for a product's price (Ellison and Wolitzky (2012)). Here, information is

[*]Ashoka University, India

[†]Federal Trade Commission, USA. The views expressed in this paper are my own and not necessarily those of the Federal Trade Commission or of any individual Commissioner.

binary and only the time it takes to acquire it is affected by the firm. Another strategy for studying obfuscation has been to focus on strategic naivete (Gabaix and Laibson (2006)). In this literature, some buyers have no or limited ability to understand the negative Bayesian implications of an obfuscated attribute, while others fully understand. What is left out, however, is the ability of these individuals to acquire additional information on their own. That is, there is no learning on the part of either type of consumer. Lastly, there is a literature on cognitive constraints in such contexts (De Clippel and Rozen (2021), de Clippel and Rozen (2022)). Here, sellers can choose to obfuscate information, causing buyers to rely on their own attention or information processing abilities rather than free information. However, they do not model the ability of sellers to precisely choose *how* such obfuscation is performed nor do they model the ability of sellers to directly impact how such information is acquired by the buyer.

In our approach, we consider a strategic communication model where buyers can obtain payoff-relevant information subject to the cognitive cost of acquiring it. This information comes in the form of a distribution of posterior beliefs. This element of our model is similar to work in Matysková (2018) and Matyskova and Montes (2023)). What we add is the ability of the seller to costlessly choose a garble of the belief distribution acquired by the buyer. The buyer then chooses an action based on the garbled information that affects the payoff of both agents. A similar approach has been studied by Linhares (2021) and Janssen and Kasinger (2022). Both papers consider strategic communication models where the buyer has an opportunity to learn by paying a *rational inattention* (Caplin and Martin (2020)) learning cost. However, in both models, instead of affecting the information of the buyer directly, the seller affects the learning cost parameter for the buyer. We will show that our model allows for a broader range of interpretations, such as hiding information, misleading information, and outright fraud. In addition, their approach does not allow us to study the heterogeneous effect of obfuscation on buyers of differing attentional abilities. Allowing the seller to garble the belief directly allows agents to have heterogeneous learning cost parameters.

Our main finding is as follows: when the buyer has a favorable opinion of the seller, i.e., would choose the preferred action of the seller given his prior, the seller optimally chooses no obfuscation in the state where preferences align but choose an intermediate level of obfuscation in the misaligned state. However, under unfavorable prior, it is optimal for the seller to choose an intermediate level of obfuscation in all states.

2 Model

Let us consider a one-shot game with two players, a Seller, and a Buyer. S holds one unit of an indivisible good that B can buy. There are two possible states of the world, $\omega \in \{\omega_0, \omega_1\}$ that denote the valuation of the good for B. The set of actions is denoted by $a \in \{a_0, a_1\}$ where a_0 denotes the action of not buying and a_1 that of buying. The payoff function is as follows:

$$u(a_0, \omega_0) = u(a_1, \omega_1) = \overline{u} > 0 \tag{1}$$

$$u(a_1, \omega_0) = u(a_0, \omega_1) = \underline{u} = 0(WLOG). \tag{2}$$

B is the Bayesian expected utility maximizer, i.e., wants to buy the good only in state ω_1. However, S always wants to sell the good irrespective of the state. We assume the prior belief of B is given by $\mu = Pr(\omega_1)$. If $\mu < 0.5$, without further information, B would choose not to buy the good. We assume S knows the state ω.

Before making the purchase decision, B can learn about the state ω subject to paying the cost of learning. We assume the cost of learning function is given by a linear function of the mutual entropy between the prior and the expected posterior belief, a la Rational Inattention (RI henceforth) literature, as follows,

$$K(\mu) = \lambda E_\omega D(P(a, \omega)||P(a)).$$

where $P_B(a)$ and $P_B(a, \omega)$ denote the unconditional (or prior) and conditional (or posterior) probability of choosing a by B respectively. Following the tradition of RI literature, we do not model the signal structure. But for simplicity, we can assume the signal space is also given by the set of actions, WLOG, i.e.,$s \in \{a_0, a_1\}$.

Moreover, we assume S can costlessly obfuscate the learning choices of B. S chooses a function $S : \{a_0, a_1\} \rightarrow \{a_0, a_1\}$ such that any belief $P_B(a, \omega)$ of B is obfuscated to a garbling of P_B as follows,

$$\frac{P_S(a, \omega)}{\mu(\omega)} = \sum_{j=0,1} \frac{P(a_j, \omega)}{\mu(\omega)} S(a_j, a) \tag{3}$$

Without loss of generality, we will assume,

$$S(a_0, a_0) = p_{00} \in [0, 1]$$
$$S(a_1, a_1) = p_{11} \in [0, 1]$$

If $p_{00} = p_{11} = 1$, then there is no obfuscation. We denote $p_{11} = 1$ and $p_{00} = 0$ as the case of full obfuscation. An intermediate level of obfuscation refers to $p_{11} \leq 1$ and $p_{00} \geq 0$ with at least one strict inequality.

The timeline of the game is as follows:

1. S chooses $S(a_j, a)$
2. B chooses $P_B(a, \omega)$ and learns accordingly
3. P_B is obfuscated to P_S
4. B chooses an action based on P_S
5. Payoffs are realized

We solve the equilibrium by backward induction. First, we solve B's learning strategy given $S(a_j, a)$. Given B's optimal learning strategy, we solve for the optimal choice of p_{00}, p_{11} by S that maximizes $P_S(a_1) = \mu P_S(a_1, \omega_1 + (1-\mu) P_S(a_1, \omega_0))$, i.e., the obfuscated unconditional probability of choosing a_1, i.e., buying the good. The following lemma describes the impact of obfuscation choice $S(a_j, a)$ on B's learning strategy.

Lemma 1. *The optimal learning strategy of B given the obfuscation choice of $S(a_j, a)$ is equivalent to the optimal learning strategy of B under distorted payoff functions as follows:*

$$v(a_i, \omega) = \sum_{j=0,1} S(a_j, a_i) u(a_j, \omega)$$

Proof. Given obfuscation $S(a_j, a)$ B's utility function becomes,

$$u(P_S(a, \omega)) = [P_S(a_0, \omega_0) + P_S(a_1, \omega_1)]\overline{u} + [P_S(a_0, \omega_1) + P_S(a_1, \omega_0)]\underline{u} \tag{4}$$

where,

$$P_S(a_i, \omega_j) = S(a_i, a_i) P_B(a_i, \omega_j) + S(a_i', a_i) P_B(a_i', \omega_j). \tag{5}$$

Plugging in the values of 5 into 4 we get,

$$u(P_S(a, \omega)) = \sum_{i,j=0}^{1} P_B(a_i, \omega_j) \sum_{j=0,1} S(a_j, a_i) u(a_j, \omega)$$

We can conclude the proof by defining $v(a_i, \omega_j) = \sum_{j=0,1} S(a_j, a_i) u(a_j, \omega)$. $\square$

The main intuition behind the lemma is since the obfuscation strategy creates a garbling of the original learning strategy of B, it is *as if* B solves a decision problem with a payoff function that generates a lower payoff in every state. The linearity of $P_S(a, \omega)$ in $S(a_j, a)$ generates the linear relationship between the two payoff functions.

Given the strategy of B, we can solve for the optimal obfuscation choice of S. Before solving the strategy choice of S let us note that if $P(\omega_1|a_1) < 0.5$, the signal a_1 is not informative about state ω_1 and thus B would choose not to learn since learning is costly. This generates the following condition,

$$Pr(\omega_1|a_1) \geq 0.5 \tag{Condition A}$$

which can be rewritten as

$$\frac{\mu P_S(a_1, \omega_1)}{\mu P_S(a_1, \omega_1) + (1-\mu) P_S(a_1, \omega_0)} \geq 0.5$$

i.e., after garbling the signals should be sufficiently informative such that the posterior probability of state ω_1 given the signal a_1 remains above 0.5. Thus the seller's problem is given by,

$$\max_{p_{00}, p_{11}} P_S(a_1)$$
$$s.t. \quad \text{Condition A}$$

Our main result in this model is outlined in the following theorem.

Theorem 1. *The following strategy profile describes the obfuscation equilibrium. The seller S choose $p_{11} = 1$ for all values of $\mu \in [0, 1]$.*

i. For $\mu \geq \frac{1}{2}$, S chooses $p_{00} = 0$, i.e., full obfuscation.

ii. For $\mu \in (\frac{1}{e^u+1}, \frac{1}{2})$, S chooses $p_{00} \in (0, 1)$, i.e., intermediate level of obfuscation.

iii. For $\mu \leq \frac{1}{e^u+1}$, S chooses $p_{00} = 1$, i.e., no obfuscation.

The buyer B,

i. Does not learn and chooses $P_S(a_1) = 1$ (or $\frac{1}{2}$) for $\mu > \frac{1}{2}$ (or $= \frac{1}{2}$)

ii. Learns according to lemma 1 and chooses $P_S(a_1) \in (0, 1)$ accordingly for $\mu \in (\frac{1}{e^u+1}, \frac{1}{2})$.

iii. Does not learn and chooses $P_S(a_1) = 0$ for $\mu \leq \frac{1}{e^u+1}$.

Proof. **Step 1: Probability of choosing** a_1 Let us begin by solving the seller's problem. For notational simplicity, let us assume $u = \frac{\bar{u}}{\lambda}$. For $p_{00} = 1$ and $p_{11} = 1$, the seller chooses not to obfuscate; the probability of choosing a_1 is given by the unconditional choice probability $P(a) \equiv p_a$ from the buyer's problem, given by

$$p_a = \begin{cases} 1 \quad \text{if} \quad \mu \geq \frac{e^u}{e^u+1} \\ \frac{e^u-(1-\mu)-\mu e^{2u}}{2e^u-e^{2u}-1} \quad \text{if} \quad \mu \in (\frac{1}{e^u+1}, \frac{e^u}{e^u+1}) \\ 0 \quad \text{if} \quad \mu \leq \frac{1}{e^u+1} \end{cases}$$

For $p_{00} = 0$ and $p_{11} = 1$, the seller chooses to obfuscate fully and the buyer always gets a_1 signal. In this case, the buyer cannot learn any further. Thus the probability of choosing a_1 is given by,

$$Pr(a_1) = \begin{cases} 1 \quad \text{if} \quad \mu > 0.5 \\ 0.5 \quad = \text{if} \quad \mu = 0.5 \\ 0 \quad = \text{if} \quad \mu < 0.5 \end{cases}$$

For any other value of p_{00} and p_{11} the seller chooses an intermediate level of obfuscation where we can rewrite $P_S(a)$ as

$$P_S(a_1) = \frac{(p_{11} + \mu(1 - p_{00} - p_{11}))e^u - (1 - p_{00} - \mu(1 - p_{00} - p_{11}))e^{u(p_{00}+p_{11})}}{e^u + e^{u(p_{00}+p_{11})}}$$

Without loss of generality, we can assume $p_{11} \geq p_{00}$, if not, relabelling the signals would generate the desired inequality.

Step 2: Relationship between $P_S(a_1)$ **and** p_{00}, p_{11} The FOC of $P_S(a_1)$ w.r.t. and p_{11} is as follows:

$$\frac{\partial P_S(a_1)}{\partial p_{11}} = (1 - \mu) + \mu e^{2u(p_{00}+p_{11}-1)} + (u(p_{00} + p_{11} - 1)(1 - 2\mu) - 1)e^{u(p_{00}+p_{11}-1)}$$

For all $\mu < 0.5$, $\frac{\partial P_S(a_1)}{\partial p_{11}} > 0$ thus optimally $p_{11} = 1$ should be chosen.

Similarly, The FOC of $P_S(a_1)$ w.r.t. and p_{00} is as follows:

$$\frac{\partial P_S(a_1)}{\partial p_{00}} = -\mu + -(1 - \mu)e^{2u(p_{00}+p_{11}-1)} + (u(p_{00} + p_{11} - 1)(1 - 2\mu) + 1)e^{u(p_{00}+p_{11}-1)}$$

For all $\mu < 0.5$, $\frac{\partial P_S(a_1)}{\partial p_{00}} < 0$ thus optimally $p_{00} = 0$ should be chosen. However, at $p_{11} = 1$ and $p_{00} = 0$, signals become fully uninformative, thus $Pr(\omega_1|a_1) = Pr(\omega_1) = \mu < 0.5$, i.e., condition A is violated.

Step 3: Relationship between $P_S(a_1)$ and p_a If $P_S(a_1) < p_a$ it is not optimal for the seller to obfuscate. We now show that if $p_{00}+p_{11} < 1$, $P_S(a_1) < p_a$. Thus under $p_{00}+p_{11} < 1$ we need to show that the following inequality holds true

$$\frac{(p_{11} + \mu(1 - p_{00} - p_{11}))e^u - (1 - p_{00} - \mu(1 - p_{00} - p_{11}))e^{u(p_{00}+p_{11})}}{e^u + e^{u(p_{00}+p_{11})}} < \frac{e^u - (1 - \mu) - \mu e^{2u}}{2e^u - e^{2u} - 1}$$

Note that, if $p_{00} + p_{11} < 1$, the denominator of LHS is positive. However, since the denominator of the RHS is always negative we can rewrite the inequality as follows:

$$\begin{aligned} &e^{2u}(2p_{11} - p_{00} - 1 + 2\mu - \mu(p_{00} + p_{11})) + e^{u(p_{00}+p_{11}+1)}(2p_{00} - p_{11} - 1 + 2\mu - \mu(p_{00} + p_{11})) \\ &+e^{2u(p_{00}+p_{11})}(1 - 2\mu - p_{00} + \mu(p_{00} + p_{11})) + e^{(3-p_{00}-p_{11})u}(1 - 2\mu - p_{11} + \mu(p_{00} + p_{11})) > 0 \end{aligned}$$

This implies,

$$\begin{aligned} 2e^{2u}(p_{11} - p_{00})(1 - e^{p_{00}+p_{11}-1}) + e^{2u}(1 - 2\mu + \mu(p_{00} + p_{11}))(e^{2u(p_{00}+p_{11}-1)} - 1)(1 - e^{u(1-p_{00}-p_{11})}) \\ +e^{2u}(e^{2u(p_{00}+p_{11}-1)} - 1)(p_{11}e^{u(1-p_{00}-p_{11})} - p_{00}) > 0 \end{aligned}$$

which is true if $p_{00}+p_{11} < 1$. Thus under obfuscation, the optimal choice of the seller would be such that $p_{00} + p_{11} \geq 1$

Step 4: $p_{11} = 1$ under obfuscation Let us prove this by contradiction. Suppose there exists an equilibrium strategy where $p_{11} = 1 - \epsilon$, where $\epsilon > 0$ and $p_{11} + p_{00} = K > 1$. Let us consider an alternate strategy of the seller such that $p'_{11} = 1 - \epsilon/2$ and $p_{00} = K - 1 + \epsilon/2$. By step 3, such a strategy would not make the no obfuscation strategy strictly better. However, by increasing p_{11} and decreasing p_{00}, the seller is strictly better off under obfuscation, contradicting the assumption that p_{11}. Hence, $p_{11} < 1$ can't happen in equilibrium.

Step 5: No obfuscation Optimal for small enough μ If $p_{00} \geq 1$ given $p_{11} = 1$ by condition A, then the seller would choose to not obfuscate. Let μ^* denote the value of μ such that $p_{00} = 1$. We can solve μ^* by solving condition A with equality for $p_{11} = 1$ and $p_{00} = 1$. Thus μ^* is such that,

$$\begin{aligned} 3e^{2u} + \mu^* e^{4u} + (1 - \mu^*) = (2\mu^* + 1)e^{3u} + (3 - 2\mu^*)e^u \\ \Rightarrow \mu^* = \frac{1}{e^u + 1} \end{aligned}$$

Step 6: Solving p_{00} for intermediate range of μ Since $p_{00} + p_{11} > 1$, Condition A

can be rewritten as,

$$3(p_{11}-(1-p_{00}))e^{2u}+(\mu p_{11}-(1-\mu)(1-p_{00}))e^{2(p_{00}+p_{11})u}+((1-\mu)p_{11}-\mu(1-p_{00}))e^{2(2-p_{00}-p_{11})u}$$
$$+((3-2\mu)(1-p_{00})-(2\mu+1)p_{11})e^{(p_{00}+p_{11}+1)u}-((3-2\mu)p_{11}-(2\mu+1)(1-p_{00}))e^{(3-p_{00}-p_{11})u}\geq 0$$

Given $p_{11}=1$ by step 4, p_{00} is determined by condition A at equality as follows,

$$3e^{2u}p_{00}+(\mu-(1-\mu)(1-p_{00}))e^{2u(p_{00}+1)}+(1-\mu-\mu(1-p_{00}))e^{2(1-p_{00})u}+$$
$$((3-2\mu)(1-p_{00})-(2\mu+1))e^{(p_{00}+2)u}-(3-2\mu-(2\mu+1)(1-p_{00}))e^{(2-p_{00})u}=0$$

We will further show that as μ decreases from $\frac{1}{2}$ to $\frac{1}{e^u+1}$, p_{00} increases from 0 to 1. Since condition A with equality generates an implicit function of μ and p_{00}, we use the implicit function theorem to obtain $\frac{\partial p_{00}}{\partial \mu}$ as follows,

$$\begin{aligned}
p'_{00}\equiv\frac{\partial p_{00}}{\partial \mu}&=-\frac{(2-p_{00})(e^{2up_{00}}-e^{-2up_{00}}-e^{up_{00}}+e^{-up_{00}})}{I+uII}\\
I&=3+(1-\mu)e^{2up_{00}}+\mu e^{-2up_{00}}-(3-2\mu)e^{up_{00}}-(2\mu+1)e^{-up_{00}}\\
II&=2(\mu-(1-\mu)(1-p_{00}))e^{2up_{00}}-2(1-\mu-\mu(1-p_{00}))e^{-2up_{00}}\\
&\quad+((3-2\mu)(1-p_{00})-(2\mu+1))e^{up_{00}}+(3-2\mu-(2\mu+1)(1-p_{00}))e^{-up_{00}}
\end{aligned}$$

The numerator can be rewritten as,

$$(2-p_{00})(e^{2up_{00}}-e^{-2up_{00}}-e^{up_{00}}+e^{-up_{00}})=(2-p_{00})(e^{up_{00}}-1)(e^{up_{00}}+e^{-2up_{00}})>0$$

if $u>0$ and $p_{00}\in[0,1]$. The first term in the denominator can be rearranged using condition A as follows,

$$I=\frac{(1-2\mu)}{p_{00}}(e^{2up_{00}}-e^{-2up_{00}}-2e^{up_{00}}+2e^{-up_{00}})=\frac{(1-2\mu)}{p_{00}}((e^{up_{00}}-1)^2-(e^{-up_{00}}-1)^2)>0$$

The second term in the denominator can be written as,

$$II=(e^{up_{00}}-1)^2(e^{up_{00}}+1)e^{-up_{00}}(e^{-up_{00}}(2(1-2\mu+\mu p_{00})(e^{up_{00}}-1))-(2-p_{00})(1-2\mu))$$

This would be positive if and only if

$$e^{-up_{00}}(2(1-2\mu+\mu p_{00})(e^{up_{00}}-1))-(2-p_{00})(1-2\mu)\geq 0$$

This can be rearranged to,

$$\Delta \equiv p_{00}e^{up_{00}} - 1 - 2\mu + \mu p_{00} \geq 0 \tag{6}$$

At $\mu = 1/2$, $p_{00} = 0$, $\Delta = 0$ and at $\mu = 1/e^u + 1$, $p_{00} = 1$, in which case, $\Delta = \frac{e^{2u}}{e^u} + 1 > 0$. This implies Δ decreases from a positive value to zero as μ increases from $1/e^u + 1$ to $1/2$. The following inequality shows the condition under which $\frac{\partial \Delta}{\partial \mu} < 0$.

$$p'_{00}(e^{up_{00}} + up_{00}e^{up_{00}}) + 4 - 2p_{00} - 2\mu p'_{00} < 0.$$

If $p'_{00} \geq 0$ for the entire range of μ, $\frac{\partial \Delta}{\partial \mu} \geq 0$. Then since at $\mu = 1/e^u + 1$ $\Delta > 0$, as μ increases Δ would increase and remain positive for $\mu \in [1/e^u + 1, 1/2]$. But if Δ increases $II > 0$ for this entire range of μ. However, this generates a contradiction since p'_{00} can only be positive if $II < 0$ or $\Delta < 0$.

If instead $\frac{\partial \Delta}{\partial \mu} < 0$ as $\mu \to 1/e^u + 1$ but becomes zero for some $\mu \in (1/e^u + 1, 1/2)$, then Δ must obtain a minima below 0 for some $\mu' \in (1/e^u + 1, 1/2)$. At the minimum,

$$\frac{\partial \Delta}{\partial \mu} = 2 - \frac{p_{00}(e^{3up_{00}+1})(e^u p_{00} - 1)((1 + up_{00})e^{up_{00}} - 2\mu)}{p_{00}I + up_{00}II} = 0$$

But at μ', since II appears in the denominator of $\frac{\partial \Delta}{\partial \mu}$, it should also attain its minima, contradicting the earlier assumption. Thus for the range of $\mu in (1/e^u + 1, 1/2)$ we find $II > 0$, which implies, $\frac{\partial p_{00}}{\partial \mu} < 0$.

Since $p_{00} = 1$ at $\mu = 1/e^u + 1$ and $p_{00} = 0$ at $\mu = 1/2$ and $\frac{\partial p_{00}}{\partial \mu} < 0$, we can conclude $p_{00} \in (0, 1)$ for all $\mu \in (1/e^u + 1, 1/2)$.

$\square$

The main intuition behind the proof is that under no obfuscation, the learning strategy of B generates more precise information than under obfuscation. If the prior belief is not in favor of S, he can improve his payoff by obfuscating. However, obfuscation cannot be so extreme that the meaning of signals is lost. This is captured by Condition A. Together we get an intermediate level of obfuscation. If however, learning was not an optimal choice for the consumer even without obfuscation, we shown that no obfuscation would be optimal. On the other hand, when the prior is favorable and S fully obfuscates he can ensure B will choose a_1 for sure, thus under favorable prior full obfuscation is optimal.

Corollary 1. *The level of obfuscation is decreasing in the marginal cognitive cost λ.*

Proof. Note that, at $\mu^* = \frac{1}{e^u+1}$, the seller chooses $p_{00} = p_{11} = 1$. But

$$\frac{\partial \mu^*}{\partial \lambda} = \frac{\partial \mu^*}{\partial u}\frac{\partial u}{\partial \lambda} = (-\frac{e^u}{e^u+1})(-\frac{\overline{u}}{\lambda^2}) > 0$$

Thus, as λ increases the seller can obfuscate for a smaller range of possible values of μ. $\square$

3 Examples

3.1 Misleading Information

Consider the following problem where an add-on is available along with a basic product. Let ω_1 denote the state where the add-on is not needed to consume the basic product and ω_0 where the add-on is needed. Examples include printers and ink, laptops and operating systems, etc. Let a_1 denote the action of buying the basic product and a_0 denote that of not buying. The state-dependent utility function is the same as described in the model (1). Thus the buyer wants to buy the basic product only if the add-on is not needed. The seller can give misleading information to the buyer in both states. Let p_{00} denote the probability that the seller truthfully reveals an add-on is needed in ω_0 and p_{11} denote the probability that the seller truthfully reveals no add-on is needed in ω_1. For the sake of simplicity let us assume, $\frac{\bar{u}}{\lambda} = 1$.

For $\mu = 0.4$, the optimal solution given by condition A is $p_{11} = 1$ and $p_{00} = 0.729$. For these values $P(a_1) = 0.43$. If the seller however chooses no obfuscation,

$$P(a_1) = \frac{\mu e + \mu - 1}{e - 1} = 0.13$$

Thus the seller is better off by obfuscating.

3.2 Hiding Information

In this example, we consider a two-dimensional good that the seller wants to sell. Suppose the value of each dimension is given by H (for high) and L (for low). Let ω_1 denote the good state where the buyer is better off by buying the good and ω_0 denote the state where not buying is optimal. Let us consider the following possibilities:

- HH and state ω_1 (HH) : 1/3
- LL and state ω_0 (LL): 1/3
- HL and state ω_1 (HL^1): 1/12

- HL and state ω_0 (HL^0): 1/4

Thus $\mu = Pr(\omega_1) = \frac{5}{12} < 0.5$. We assume in this problem the seller can hide information about either dimension and reveal the value (H or L) for only the other dimension. Consider the following obfuscation strategy:

- HH: report H
- HL^G: report H w.p. q
- HL^B: report H w.p. q'
- LL: report L

The one-to-one relationship between q, q' and p_{ii} is as follows,

$$p_{11} = Pr(H|HH \cup HL^G) = \frac{4+q}{5}, \quad p_{00} = Pr(L|LL \cup HL^B) = \frac{4+3(1-q')}{7}$$

In this problem, the optimal solution is given by $p_{11} = 1$ and $p_{00} = \frac{2}{3}$. Thus, $q = 1$ and

$$\frac{4+3(1-q')}{7} = \frac{2}{3} \Rightarrow q' = \frac{7}{9}$$

denote the optimal strategy. In this case, also, the seller has a higher probability of selling by hiding information.

4 Conclusion

In this paper, we build a general model of obfuscation behavior. We find that an intermediate level of obfuscation is optimal under unfavorable prior and full obfuscation is optimal only under favorable prior. Moreover, we show that our framework can be used to describe different forms of obfuscation behavior namely, misleading advertisement (example 2), direct fraud (example 1), etc.

We have preliminary evidence for our main theoretical result (Theorem 1). Fehr and Wu (2023) showed in an experimental market of a base good with add-on features that sellers choose to obfuscate only when add-on features are surplus enhancing. However, the obfuscation level diminishes significantly if add-on features are surplus neutral. However, their setup did not capture prior beliefs directly. Moving forward we want to construct a market experiment to directly test our results.

Bibliography

Caplin, A. and D. J. Martin (2020). Framing, information, and welfare. Technical report, National Bureau of Economic Research.

De Clippel, G. and K. Rozen (2021). Bounded rationality and limited data sets. *Theoretical Economics 16*(2), 359–380.

de Clippel, G. and K. Rozen (2022). Communication, Perception and Strategic Obfuscation. *Working Paper*.

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

Fehr, E. and K. Wu (2023). Obfuscation in competitive markets. Technical report, Working Paper.

Gabaix, X. and D. Laibson (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics 121*(2), 505–540.

Janssen, A. and J. Kasinger (2022). Obfuscation and rational inattention. Technical report, Working Paper.

Linhares, L. H. R. (2021). *Persuasion by strategic obfuscation*. Ph. D. thesis.

Matysková, L. (2018). Essays on information economics.

Matyskova, L. and A. Montes (2023). Bayesian persuasion with costly information acquisition. *Journal of Economic Theory*, 105678.