ABSTRACT

Title of Dissertation: ESSAYS ON RATIONAL INATTENTION IN INDIVIDUAL AND STRATEGIC DECISION MAKING

Erika Dömötör
Doctor of Philosophy, 2022

Dissertation Directed by: Professor Erkut Ozbay
Department of Economics

Chapter 1 revisits the model of adverse selection under asymmetric information with the power of the rational inattention framework. I depart from the setup of Akerlof (1970) by revising its extreme information asymmetry assumption. Instead of assuming that the Seller is fully informed and the Buyer is fully uninformed, I consider a setting in which both parties are able to gather information, but at a cost. As a result, both the Seller and the Buyer become partially informed, and the information asymmetry is the consequence of the asymmetry in their incentives and unit information costs. This enhanced framework provides new insights into the implications of incomplete information for market outcomes, efficiency and welfare. When information asymmetries occur endogenously, they do not lead to market collapse, but they do create market inefficiencies. The Buyer is better off and the Seller is worse off compared to the efficient symmetric information benchmark.

In Chapter 2, I propose a model that explains the evolution of overconfidence as being a result of the constrained utility-maximizing problem of a decision maker who is rationally inattentive to information, but at the same time biased towards more optimistic subjective beliefs.

Empirical studies have shown that individuals with initially fewer skills have more confidence, but as their skill level increases, their overconfidence decreases. The phenomenon is well-known as the Dunning-Kruger effect in the psychology literature. I explain this effect by the simultaneous choice of subjective and objective information. In my model a non-materialistic utility component induces overly optimistic subjective beliefs, which are constrained by the cost of information distortion. The setup is tractable in a range of economic problems.

Chapter 3 utilizes the Model of Overconfidence from Chapter 2 in an application which explains the excess entry and high drop out rate of entrepreneurs. In this setting I show that in the presence of overconfidence individuals enter businesses with lower than necessary skill levels to succeed. At the same time, they drop out due to underperforming even when their skill levels would be adequate to stay in, were they not overconfident. The gap between skill thresholds for entry and drop out results in the high failure rate of businesses.

ESSAYS ON RATIONAL INATTENTION IN INDIVIDUAL AND STRATEGIC DECISION MAKING

by

Erika Dömötör

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:
Professor Erkut Ozbay, Chair
Professor Emel Filiz-Ozbay
Professor Yusufcan Masatlioglu
Professor Luminita Stevens
Professor Emanuel Zur

Dedication

To my Family

Acknowledgments

I would like to express my gratitude to all who contributed towards the production of this dissertation one way or another. The positive change you made in my career and in my personal life is tremendous.

I thank first my main advisor, Professor Erkut Ozbay. He was involved in every steps of a long process, and provided invaluable support. He should be thanked for each and every of these moments, however here let me highlight only the two I appreciated the most. I am grateful for his inspiration and guidance in the early stages of identifying and formalizing the research topics which keep me excited even beyond these chapters. I am also deeply thankful for his patience and optimism he exhibited whenever I faced obstacles. His positive attitude helped me to stay on track, and eventually to grow beyond my limits.

I would like to thank all members of my dissertation committee. I thank Professor Luminita Stevens for introducing me into the techniques I apply in Chapter 1. I am also grateful for her broader perspective from outside of my narrower field. I would like to thank Professor Emel Filiz-Ozbay and Professor Yusufcan Masatlioglu for teaching me to see the difference between good and not so good theories. I am also grateful to both of them for sharing some of their personal experiences as graduate students and later as academics, which further encouraged me to pursue research. I am also indebted to Professor Emanuel Zur, for his time and willingness to serve in my committee.

I thank all other faculty and staff members, and fellow graduate students of the Department of Economics. I am especially grateful to Professor Allan Drazen and Professor Daniel Vincent for their consistent presence and feedbacks during my seminar talks. I am thankful to former and current graduate students in the Behavioral and Experimental Economics group for the discussions through several official and unofficial meetings. I would like to thank the staff of the main office for their work and day-to-day kindness. Finally, I am utterly grateful to Professor John Shea, for being the sentinel over all of my endeavors in graduate school, who was always there to smooth out any academic or non-academic tangles.

I would like to express my gratitude to my parents who raised me to have the courage and curiosity to leave the comfort and explore unknown territories. I thank two wonderful sisters who followed my adventures from the distance, and cheered all of my experiences and achievements. I consider myself blessed to meet exceptional new friends along the way. I would like to especially thank Dániel Baksa, Natalia Bogdanova, Maria Krutkova and Béla Személy for the countless shared rides, meals and conversations.

I am beyond words grateful to two incredible communities, Carpathia Folk Dance Ensemble and Tisza Ensemble. I thank them for their pure existence, and everything they stand for. They were the bridge through which I could finally arrive in a new world, the soil where I could start growing roots. I grateful to all of their members for lending me a bit of their energies when I needed to recharge the most. I thank them for their acceptance and for becoming my family away from home.

My final word of gratitude I would like to dedicate to the person who was right on my side through the major part of my time in graduate school, and who became one true ally in all the ups and downs. I thank him for the inspiration, for the challenge, and for the care.

Table of Contents

List of Tables

List of Figures

Chapter 1: Bringing Rational Inattention to the Market for Lemons

1.1 Introduction

In neoclassical economic models the information that is available and used for decision making is specified with absolute clarity. However, real world decisions are encompassed by the choice of the process of decision making itself; including the free choice of gathering or ignoring information before any action is taken. For example, a consumer browsing an online shopping website has access to a large amount of data about the products' characteristics and values, but she may not use all of it to come to a decision about her purchase. My project accounts for the complexity of the utilization of information in a framework where attention is chosen based on its benefits and costs. For an individual decision maker, the benefit is the increased chance of taking a more gainful action, and the cost is the cognitive effort associated with paying attention. In the online shopping example, the gain from being more informed is higher when the stakes are higher, such as when a costly electronic device is being purchased. In this case the consumer would be willing to pay more attention in order to increase her chances to make the optimal decision. The cost of attention can be lower or higher depending on how accessible and straightforward the provided information is. For example, a more organized, user-friendly website can lower the attention costs to the consumer, which encourages her to gather more information before the purchase.

In strategic situations individuals choose both their actions and information in a way that their choices best reply to others' decisions, and to other elements of the environment. The costs and benefits of attention may greatly vary between individuals, which gives the edge to one or another agent in the information gathering process. Consequently, information asymmetries may arise, and they can occur in many different structures and magnitudes depending on the given situation. For example, on durable good markets, such as the used-car market, the Seller, the current owner and user of the vehicle, knows more about the quality of the good than the Buyer. However, in a different situation the Buyer may be the one who is more knowledgeable about the value of the product. For example, a Buyer who is a collector looking for antiques at a yard sale is more informed about the good than the Seller himself.

Information asymmetries can entail major impacts on economic decisions and outcomes. In Akerlof's (1970) "market for lemons" model a Seller who knows everything about his product interacts with a Buyer who is, contrariwise, uninformed about the product's actual quality and relies only on commonly known prior information. In his setting information asymmetry leads to an adverse selection problem, which leaves only low quality goods – the so called "lemons" – on the market. The Seller, knowing the true quality, only sells his good when its value is low, while the Buyer, knowing that only low quality goods are sold, rejects all offers. Thus, no trade occurs and the market shuts down completely. In an analogous setting with the only modification that the information to the agents' are symmetric the market would not collapse. Therefore, adverse selection emerges neither due to the lack of mutual interest in trade, nor due to the lack of information, but due to the asymmetry in information.

Adverse selection is relevant in several economic contexts, for example, in health insurance markets [Culter and Reber, 1998, Cardon and Hendel, 2001, Powell and Goldman, 2021], in decentralized markets, such as in labor markets [Barsanetti and Camargo, 2021] and in durable good markets, such as the used car or housing markets [Hendel and Lizzeri, 1999]. Adverse selection was also named as one of the possible drivers of financial crises. Tirole [2012], and Philippon and Skreta [2012] discuss possible interventions to prevent market failure by removing the "lemons" from the asset markets. While these interventions successfully restore the operation of the markets, they are generally costly. In contrast, my study, which accounts for the free choice of information, offers an alternative approach to make those tolls lower. Information asymmetries can be mitigated by lowering the attention costs for one or both agents, for example, by information disclosure requirements or by centralized data bases. These interventions specifically target the information asymmetry problem, and reveal the "lemons" rather than removing them from the market.

In this paper I propose a generalized framework of information asymmetry, and apply it to analyze the extent of market inefficiencies due to adverse selection and the effect of possible interventions through the costs associated with attention. I allow for the possibility that even the most experienced Seller faces some level of uncertainty, and even the most naive Buyer might have a chance to inspect the good and gather some additional information about its value. In this generalized framework information is the choice of the agents; both the Seller and the Buyer choose a reduced level of uncertainty based on the expected benefits and costs. This modified setup provides more flexibility to describe the role of information in the adverse selection problem while still captures asymmetry arising from the different costs per unit of information and from the different benefits from reducing uncertainty. Akerlof [1970] predicts that information asymmetry generates a market collapse. In contrast, my model implies that the market does not collapse, every quality of good is sold with positive probability; however, they are never sold with certainty.

In the current setup, I replace the extreme information asymmetry with a more general assumption, in which information is the endogenous choice of each agent. More specifically, both the Seller and the Buyer, facing a common prior distribution over the value of the good, can choose a set of signals that modify their prior beliefs to a more precise posterior distribution. I utilize the rational inattention framework to describe the agents' choices and quantify the information content of the signals. [Sims, 2003a] This setup does not pre-specify the format of the information, hence each agent chooses an information structure – a set of signals – which serves their individual decisions best. The only restriction is that the chosen information structure has to be consistent with the prior information, and that the signal sets are subject to information costs. Woodford [2009] shows that the agents optimally choose signals that are tailored to their feasible actions, hence they can be seen as direct recommendations about which action should be taken. That means a signal for the Seller is a price he should offer to the Buyer, and the signal for the Buyer is the decision whether or not she should accept this offer.

The cost of an information structure is defined based on the uncertainty reduction that the included signals generate, where the uncertainty is measured in the entropy of probability distributions. [Shannon, 1959] This information measure has two advantages in my setup. First, information costs are quantified in functional forms, which are straightforward to implement in utility-maximization problems. Second, these cost functions are defined based on the posterior distributions generated by the signals. That means the information level is measurable and comparable regardless of the form of the signals each agent chooses.

The setup of my model consists of four-stages. In the first stage the Seller chooses his information structure. In the second stage, Nature draws the value of the good and sends a signal to the Seller from his chosen information structure, who then chooses the price of the product. In the third stage, after observing the price offer from the Seller, the Buyer updates her beliefs and chooses her information structure. Finally, in the fourth stage, the Buyer gets a signal from her chosen information structure and chooses either to accept or to reject the offer.

If the offer is accepted, then the trade takes place at the price the Seller quoted, and the agents collect gains according to the their valuations and the price. If the offer is rejected, then the trade does not take place, and both parties gain nothing.

The information asymmetry between the agents is represented by individual-specific values for the unit costs of information. For example, following the intuition behind the adverse selection model that the information asymmetry favors the Seller, one can suppose that the unit cost is lower for the Seller, who hence gathers information more easily than the Buyer. In the used car example the Seller might know more about the quality than the Buyer, but still needs to do research in order to estimate the optimal market price. Therefore, the Seller's advantage of knowing the good's quality better is reflected in the lower costs associated with a more accurate assessment. Similarly, in the collector of antiques example the Buyer might be more informed about market prices than the Seller, but she still might need to inspect the good for an accurate valuation. In this case the information advantage of the Buyer can be reflected in her lower unit information cost. Finally, the rational inattention approach allows for taking into account not just the effect of the difference in the costs of information, but also the effects of incentives for gathering more accurate information. When the unit information costs are kept equal, then the agent who has more incentive for accuracy acquires more information.

My model suggests that information asymmetry advancing the Seller, which can lead to adverse selection problem, arises endogenously when the Seller has sufficiently lower unit information cost than the Buyer. The generalized framework does not predict a market collapse, however it predicts some inefficiencies. I find that any quality of good is sold with positive probability, but never with certainty. The probability of successful selling is increasing in the quality, therefore the average quality of the sold items is higher than the overall average quality.

In terms of the gains from the trade, the overall profit is positive, but it is below the efficient level which is achieved in the symmetric information benchmark. In this benchmark case all profits are extracted by the Seller due to his first mover advantage. In contrast, my model predicts that both agents get a positive share from the gains of the trade. Therefore the Seller is worse off, but the Buyer is better off than in the symmetric information case. Finally, comparing the results to the predictions of Akerlof [1970], I find that both the Seller and the Buyer are better off under the endogenous information choice assumption, because the trade takes place with positive probability.

I analyze the effects of changes in the unit information cost parameters on welfare. I show that when the Buyer's unit information cost is fixed, and the Seller's information cost decreases then both the Seller and the Buyer benefit from the Seller's information cost advantage. However, the gains from the additional information mostly increases the Seller's profit. Moreover, when the amount of information is fixed, but the information gap increases because of the decreasing unit information cost of the Seller and increasing unit information cost of the Buyer, then the Seller slightly benefits and the Buyer loses from the profits. The overall profit also decreases in this case due to the increasing information asymmetry.

The paper is structured as follows. Section 2 reviews previous literature. Section 3 describes the setup of the model. Section 4 solves the model for a wide range of unit information cost parameters. Section 5 discusses the implications on market outcomes and welfare consequences.

1.2 Literature Review

The problem of adverse selection was first introduced in Akerlof's seminal paper (1970). He proposed a model of the "markets for lemons," in which the information asymmetry results in the failure of the market. Akerlof's theoretical argument was prompted an extensive empirical investigation into the existence and prevalence of the “market for lemons”. Hoffer and Pratt [1987] summarize this scholarly dialogue. They report that Bond [1982, 1984] looked at a data set of used pick-up trucks and conclude that there is no evidence of market inefficiencies. Pratt and Hoffer [1984, 1985] found empirical evidence in the same data that this market suffers from adverse selection. Metzger [1983] contributed to the argument with a theoretical model of the effect of mandatory inspection of used vehicles. He found that this law only reduces adverse selection if the cost of inspection is sufficiently low. Following this, Pratt and Hoffer [1986] looked at the used pick-up truck data again, and found supporting evidence of adverse selection due to unsuccessful interventions.

Adverse selection is relevant in several economic contexts, for example in durable good markets, such as the used car or housing markets [Hendel and Lizzeri, 1999], or in health insurance markets [Culter and Reber, 1998, Cardon and Hendel, 2001, Powell and Goldman, 2021]. Some recent projects extend the adverse selection model with additional features of the environment. For example, Barsanetti and Camargo [2021] present a model with matching frictions and adverse selection on decentralized markets, such as the labor market. Siegenthaler [2017] tests experimentally whether cheap talk can mitigate inefficiency due to information asymmetry. Philippon and Skreta [2012] and Tirole [2012] discuss the optimal government intervention to rejuvenate financial markets after bank crises.

Closer to my work, a couple of projects implement rational inattention into applications that involve strategic interactions. Martin [2017] presents a model where a fully informed Seller and an initially uninformed Buyer, who has the chance to gather information at a cost, interact in an ultimatum bargaining game. The quality of the good can be either high or low. He finds that the high quality good is sold at high price, while the Sellers with low quality goods randomize over offering a low price or mimicking the high price. The corresponding experiment in Martin

[2016] finds evidence of the above described equilibrium strategy. Moreover, the experimental results suggest that not only the player who has access to the costly information source acts as if she chose based on a rational inattention framework, but also her opponent responds as if he understood and took into account the costly information acquisition process. Similarly, another project by Ravid [2020] presents an ultimatum bargaining setup, where the fully informed Seller makes a take it or leave it offer to the uninformed Buyer who has access to costly information. The focus of his paper is to explore all possible perfect Bayesian Nash equilibria and select those which do not contain incredible information threats.

My project is different from these previous works in two ways. First, I assume that both the Seller and Buyer, and not just the Seller, choose costly information endogenously. Second, I present a problem where the action space for the Seller, the set of possible prices, is continuous. Solving the information problem in the rational inattention framework with continuous action space requires the exploitation of different techniques that did not appear in the above mentioned works.

Jung et al. [2019] mark out a methodology to solve the costly information acquisition problem with continuous action space and provide the general properties of the solution. Their main finding is that even though the action space is continuous, it is optimal to concentrate attention only on distinguishing between a discrete subset of all possible actions. This work is a generalization and economic application of the previous result of Fix [1978] in information technology. Compared to this previous paper, Jung et al. [2019] cover a more general form of the objective function and extend the theory to multivariate cases. Moreover, the notations are tailored to economic applications, which are demonstrated on a multivariate portfolio choice problem.

There are already a handful of applications in macroeconomics of this newly introduced technique. Matějka [2015] shows that price rigidity can occur when a Seller is rationally inattentive, even without adjustment costs. Stevens [2019] develops a model that explains coarse pricing policies by rational inattention. The latter paper uses the Blahut-Arimoto algorithm to generate results under different parameter assumptions. [Blahut, 1972, Arimoto, 1972, Rose, 1994] My project contributes to this growing body of literature by another application. To the best of my knowledge, this work is the first that implements the rational inattention framework with continuous action space into strategic setting.

In a broader sense, my project belongs to a set of papers investigating the role of information in the adverse selection problem. Levin [2001] revisits the model of the "market for lemons" and asks whether greater asymmetries reduces the efficiency of trade. The answer is not unambiguously yes, because while better information for the Seller lowers the demand by worsening the Buyer's curse it may shift the Supply as well. However, improving the Buyer's information, for example by making some of the Seller's private information public, has generally positive effect on trade efficiency. Similarly, Kessler [2001] observes the non-monotonicity of trade in information. Compared to these previous works, my project utilizes a more general measure of information. Levin's (2001) model uses partitioned information, while Kessler [2001] uses beliefs that are stochastically ordered. Neither of these assumptions, or any other restriction of the format of information is needed in my approach.

A couple of papers focus on the Buyer's optimal choice when the Seller is being uninformed. Roesler and Szentes [2017] explore the least informative Buyer optimal signal structure and its effect on market efficiency and welfare. They find that the trade always occurs, and characterize a range of different consumer-producer surplus pairs, while the price and outcome remains the same across all optimal signal structures. They also show that the Buyer is better off, but the Seller is worse off compared to the full information benchmark, which result is in line with my conclusion. Ravid et al. [2019] extend the framework from free information to cheap information. They compare the consequences of having information to having access to cheap information. They find that as cost of information approaches zero, the trade equilibrium goes to the the worst equilibrium under free information. Their message is that while cheap information may increase welfare, having free information may be much better.

1.3 Model

In Akerlof's (1970) "market for lemons" model of adverse selection by the distribution of the information is fully asymmetric; the Seller knows the exact value of the good, while the Buyer only knows its prior distribution. Formally, the setup I depart from is the following.

Setup (Akerlof's Market for Lemons)**.** Denote the value of a good by v, and its prior distribution by $v \sim g(v)$. After Nature draws the value of v, the Seller learns this value, while the Buyer relies on the prior information $v \sim g(v)$. Then the Seller and the Buyer interact in a two-stage game. In the first stage, the Seller offers price p to the Buyer. In the second stage the Buyer chooses to accept or to reject the offer. The net payoffs are $(p - U_S(v))$ for the Seller and $(U_B(v) - p)$ for the Buyer if the offer is accepted, and 0 for both if the offer is rejected.

Using utility functions $U_S(v) = v$, $U_B(v) = \frac{3}{2}v$ and prior distribution $v \sim U(0,1)$ one can show that no trade occurs in this market. By observing price p, the Buyer's conditional expected valuation becomes $\frac{3}{2}\frac{p}{2} = \frac{3}{4}p$, which is lower than price p. In contrast, since the Buyer's utility from the good is higher than the Seller's utility for any v, the trade would occur either if both parties had full information about the value of the good, or both relied on the common prior.

1.3.1 Information Choice of the Seller

In the timing of the game, the Seller chooses his information first. The Seller faces the prior distribution $g(v)$ and chooses an information structure which sends him a signal from a set of possible signals. The signal suggests the optimal price that the Seller should quote. Formally, the Seller's information choice is the joint distribution $f_S(p,v)$ over value v (the state of the world) and the price signal p (the action). Given this choice, Nature draws v from $g(v)$ and sends a signal from the conditional distribution $f_S(p|v)$. The Seller always optimally offers the price suggested by the price signal. [Woodford, 2009]

In the current setup both the support of the distribution over value v and the support of the distribution over price signal p are continuous. More precisely, the Seller's information is quantified as the uncertainty reduction between the prior distribution $g(v)$ and the joint posterior distribution $f_S(p,v)$:

$$\begin{aligned} I_S(g(v), f_S(p,v)) = h(g) - h(f_S) = \\ &= \int_v \int_p log(f_S(p,v)) f_S(p,v) dp dv - \\ &- \int_v \int_p log\left(g(v) \int_{v'} f_S(p,v') dv'\right) f_S(p,v) dp dv \end{aligned} \tag{1.1}$$

where $h(f(x)) = -\int_x f(x) \log(f(x))$ is the entropy of the continuous distribution $f(x)$. [Shannon, 1959]

The existing literature on the rational inattention framework with continuous action spaces suggests that the optimal signal set contains only a finite number of signals with nonzero probability weights. [Jung et al., 2019] As I show it later, this is indeed the case in the current setup. More formally, I denote the discrete set of possible prices which are chosen with positive probability by $\mathcal{P}$. Therefore, the Seller's information choice is the set $\mathcal{P}$ and the joint probability distribution $f_S(p,v)$ over the state variable v and the randomly drawn signal $p \in \mathcal{P}$ such that $\sum_{p\in\mathcal{P}} f_S(p,v) = g(v)$ for all v. I denote the posterior conditional distribution of the Seller after getting price signal p by $q_S(v|p)$. It describes the Seller's knowledge about value v after collecting information.

Alternatively, the Buyer's information choice can be described by the conditional posterior distribution over his decision p for a given v. Formally, the conditional posterior distribution $f_S(p|v)$ is connected to the posterior joint distribution through the equation

$$f_S(p,v) = f_S(p|v) \cdot g(v) \tag{1.2}$$

where $f_S(p|v)$ is the probability that the Seller offers price p when the unobserved value is v.

The cost of information for the Seller is C_S, which is linear in the entropy reduction by the observed signal:

$$C_S = \lambda_S * I_S(g, f_S) \tag{1.3}$$

The parameter λ_S is the unit cost of information of the Seller, while $I_S(g, f_S)$ represents the amount of information the Seller acquires.

1.3.2 Information Choice of the Buyer

When the Seller makes a price offer, it carries information about the value of the good that the Buyer can perceive. I assume that she fully understands this information, and modifies her prior beliefs to the posterior conditional beliefs of the Seller. After getting the price offer and forming her prior beliefs, the Buyer has two options, she either accepts or rejects the offer. Therefore, when she gathers her own information, she focuses only whether or not the offer is worth to be accepted. Therefore her information choice is the joint distribution over value v (state of the world) and actions Accept (A) or Reject (R).

Formally, after observing price p, the Buyer updates her prior beliefs to $q_B^p(v) \equiv q_S(v|p)$ and chooses the joint distribution $f_B^p(D,v)$, where $D \in \{A,R\}$. The Buyer's information is quantified as the uncertainty reduction between the prior distribution $q_S(v|p)$ and the joint posterior distribution $f_B^p(D,v)$, where the uncertainty is measured by the entropy of continuous distributions:

$$
\begin{aligned}
I_B(q_B^p, f_B) = h(q_B^p) - h(f_B) = \\
&= \sum_{D\in(A,R)} \int_v log(f_B^p(D,v))f_B^p(D,v)dv - \\
&- \sum_{D\in(A,R)} \int_v log\left(q_S(v|p)\int_{v'} f_B^p(D,v')dv'\right) f_B^p(D,v)dv
\end{aligned}
\tag{1.4}
$$

where h is the entropy of continuous distribution, as in the Seller's information choice.

Note that the information choice of the Buyer takes place after the Seller announces the price, therefore it depends on p. For consistency, the marginals of the joint distribution chosen by the Buyer have to match her prior distribution: $\sum_{D\in\{A,R\}} f_B^p(D,v) = q_S(v|p)$ for all p and v. I denote the Buyer's posterior conditional distribution over the value after observing price p and getting signal D by $q_B^p(v|D)$. It describes the Buyer's knowledge about value v after inferring from the Seller's offer and collecting her own information.

Alternatively, the Buyer's information choice can be described by the conditional posterior distribution over her decision D for a given v, after observing price p. Formally, the conditional posterior distribution $f_B^p(D|v)$ is connected to the posterior joint distribution through the equation

$$
f_B^p(D,v) = f_B^p(D|v) \cdot q_B^p(v) \tag{1.5}
$$

where $f_B^p(A|v)$ is the probability that the Buyer accepts price p when the unobserved value is v, and $f_B^p(R|v) = 1 - f_B^p(A|v)$ is the probability that she rejects the offer.

The cost of information for the Buyer is C_B^p, which is linear in the entropy reduction by the observed signal:

$$C_B^p = \lambda_B * I_B(q_S^p, f_B^p) \tag{1.6}$$

The parameter λ_B is the unit cost of information of the Buyer, while $I_B(g, f_S)$ represents the amount of information the Buyer acquires.

1.3.3 Setup and Timing of the Game

In the current seting, smilarly to Akerlof's Market for Lemons setup, the interaction of the Seller and the Buyer is described as a sequential game, but it is extended by stages where the information is endogenously chosen by either the Seller or the Buyer. In this setup I assume that by offering a price the Seller reveals his posterior belief about the valuation of the good. This assumption can be verified as long as the solution is a separating equilibrium.

Setup (Market for Lemons with Endogenous Information)**.** Denote the value of a good by v, and its prior distribution by $v \sim g(v)$, which is initially common knowledge to the Seller and the Buyer. Then the Seller and the Buyer interact in a four-stage game. The timing of the game is the following:

1. The Seller chooses an information structure; a discrete set of signals $\mathcal{P}$, and joint distribution $f_S(p, v)$ over the state variable v and signals $p \in \mathcal{P}$, such that $f_S(p, v)$ is consistent with the prior distribution $g(v)$.

2. Nature draws signal p from $\mathcal{P}$ according to the joint distribution $f_S(p, v)$ and the true state of the world v. The Seller observes the signal and offers price p.

3. The Buyer observes price p, and chooses information structure $f_B^p(D,v)$ over the state variable v and signals $D \in \{A,R\}$

4. Nature draws signal $D = A$ or $D = R$ according to the joint distribution $f_B^p(D,v)$ and the true state of the world v. The Buyer observes the signal, and Accepts or Rejects the offer, accordingly.

The net payoffs are $(p - U_S(v))$ for the Seller and $(U_B(v) - p)$ for the Buyer if the offer is accepted, and 0 for both if the offer is rejected.

The setup described above is a finite sequential game which can be solved by backwards induction. The two steps of main interest in this project are Stage 1 and Stage 3, where the Seller and the Buyer choose their information structures. In Stage 2 and Stage 4 the Seller and the Buyer always follow their signals. It is by the design of the rational inattention framework, and it easy to verify that this is optimal to do so. [Woodford, 2009, Stevens, 2019]

1.4 Characterizing Interior Solutions

I call a solution interior if after observing a price the Buyer gathers additional information, that is, if her optimal information structure puts positive weight on both accepting and rejecting the offer, for any price offered in equilibrium. I show that the interior solution exists for a range of parameters where both the Buyer's unit cost of information parameter and the asymmetry between the Seller's and Buyer's unit cost parameters are sufficiently low.

1.4.1 The Buyer's Optimization Problem

By backwards induction, I solve the Buyer's problem first. The Buyer maximizes her expected profit from choosing her information structure net of information costs.

Problem (The Buyer's Information Choice)**.**

$$\max_{f_B^p} EU_B^p - C_B^p \tag{1.7}$$

$$s.t. \quad q_S(v|p) = f_B^p(A,v) + f_B^p(R,v) \quad \forall v \tag{1.8}$$

$$f_B^p(D,v) \geq 0 \quad \forall D, v \tag{1.9}$$

where

$$EU_B^p = \int_v U_B(p,v) f_B^p(A,v) dv \tag{1.10}$$

is the expected utility of the Buyer, and

$$C_B^p = \lambda_B * I_B(q_S^p, f_B^p) \tag{1.11}$$

is the cost of information for the Buyer.

Equation 1.7 is the objective function of the Buyer. The first term is her a priori expected utility from the chosen information structure. She earns $U_B(p,v)$ if she accepts the offer, which she does so with probability $f_B^p(A,v)$, and she 0 if she rejects the offer. The second term is the cost of information, which was introduced in Section 3.2.

Proposition 1. The solution to the Buyer's information problem is

$$f_B^p(A,v) = q_S(v|p) \frac{1}{1 + exp\left(-\frac{1}{\lambda_B} U_B(p,v)\right)} \tag{1.12}$$

where $f_B^p(A,v)$ and $f_B^p(R,v) = 1 - f_B^p(A,v)$ describes the joint distribution over the Buyer's action $D \in \{A, R\}$ and the state variable v.

For the more intuitive interpretation, one can derive the probability of acceptance conditional on value v:

$$f_B^p(A|v) = \frac{1}{1+exp\left(-\frac{1}{\lambda_B}U_B(p,v)\right)} \tag{1.13}$$

and, similarly, the conditional probability of rejection is $f_B^p(R|v) = 1 - f_B^p(A|v)$.

The probability of acceptance depends on the price p through the Buyer's utility $U_B(p,v)$, and on the Buyer's unit information cost parameter λ_B. On Figure 2.1 I illustrate the acceptance probabilities as functions of the Buyer's utility for a set of unit information cost parameters. Intuitively, when the Buyer is able to gather costly information, then she accepts the offer with higher probability when the value is higher. As the cost for information decreases, the Buyer is able to gather more accurate information. It is seen on the graph that for lower λ_B parameters, represented by the orange and red lines, the probability of acceptance remains low for low values, then steeply increases and stays high for higher values. When $\lambda_B \to \infty$ then the probability of acceptance is a step function; it is 0 when the utility for the consumer is below 0, and 1 when the utility for the consumer is above zero. This means that when the cost of information approaches zero, the Buyer is able to fully distinguish between profitable and not profitable offers.

Further on, in Example 1, I illustrate the result of the Buyer's decision problem with specific parameters and utility function.

Example 1. Assume that the Buyer's prior belief is the uniform distribution over the unit interval: $q_S(v|p) = 1$ for $v \in [0,1]$. The Buyer's utility is $U_B = 3/2v - p$. The unit cost of information for the Buyer is $\lambda_B = 0.1$, and the price offer is $p = 1$

Note that without gathering additional information, the Buyer would not accept this price offer. This is because in this example the prior distribution of the Buyer after observing price p is $q_B^1(v) \sim U(0,1)$. Then the Buyer's expected gain from accepting the offer is $\frac{3}{2}E(v) = \frac{3}{4}$,

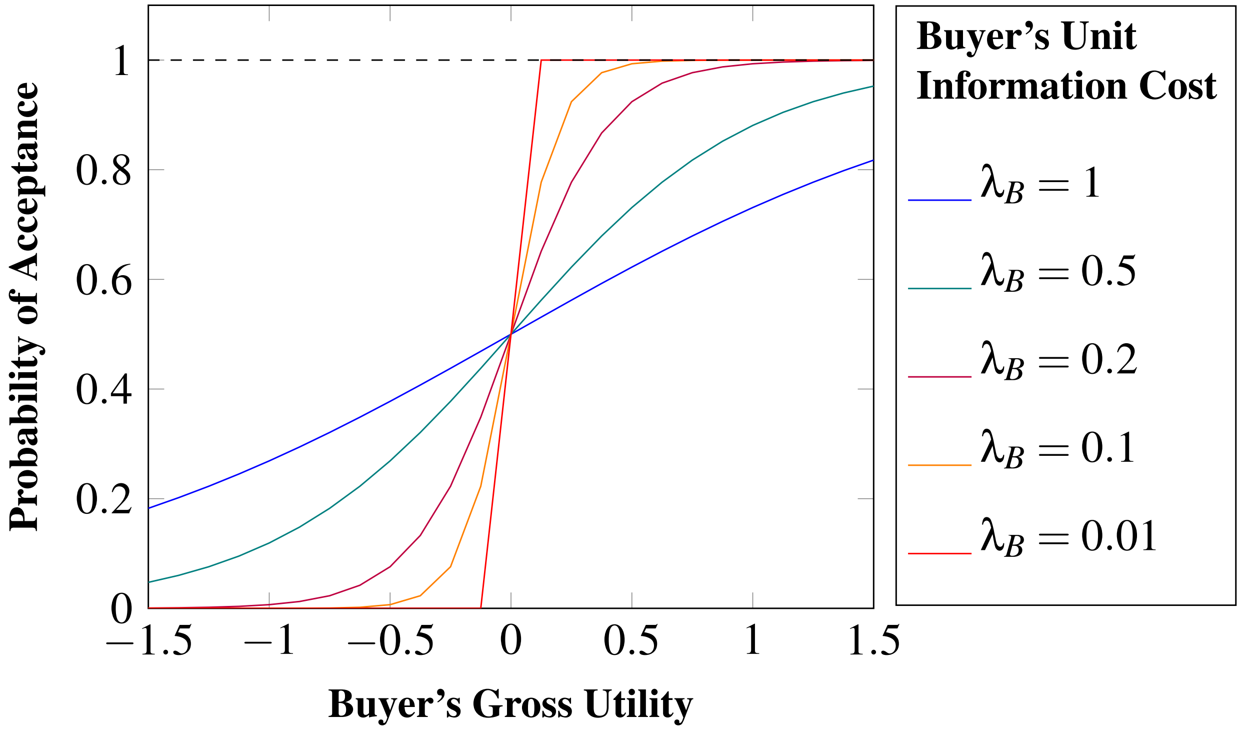


Figure 1.1: Probability of Acceptance for Various Unit Information Costs

which is below the price $p = 1$. However, when the setting allows for information gathering, the Buyer's probability of acceptance becomes $f_B^1(A|v) = \frac{1}{1+exp\left(-\frac{1}{0.1}\left(\frac{3}{2}v+1\right)\right)}$.

Figure 2.2 depicts the results of Example 1. The blue line is the prior distribution of the Buyer $q_B^1(v) \sim U(0,1)$. The red curve represents the probability of acceptance as a function of the value $f_B^1(A|v)$. The graph suggests that the offer $p = 1$ is accepted with positive probability for every v. The magnitude of this probability depends on the exact value, it is higher when v is higher. The red shaded area under the probability of acceptance curve represents the size of the active market, that is the unconditional probability that the good is sold at price 1. In this example it is 33%, which means that the current model, which allows the Buyer to get partially informed, suggests that the market does not break down completely, but it operates with some inefficiency. Therefore, the availability of costly information for the Buyer increases the probability of successful trade. The relatively low probability is driven by both the relatively high price, and the fact that the Buyer's information strategy necessarily involves putting positive

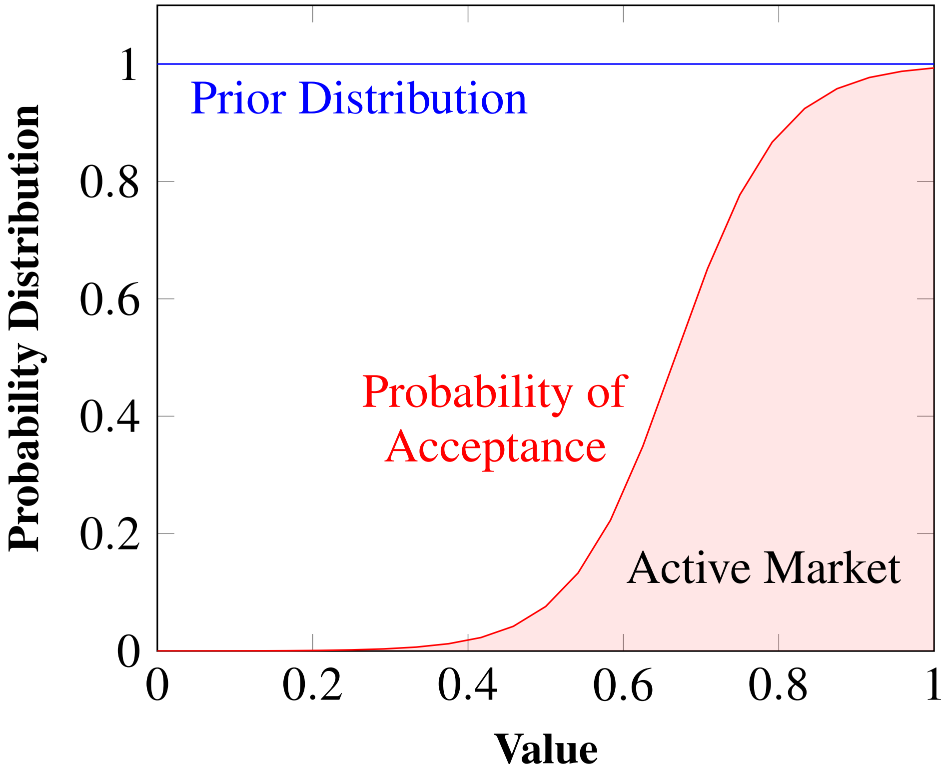


Figure 1.2: The Buyer's Optimal Information Structure ($\lambda_B = 0.1$, $p = 1$)

probability on rejecting the offer.

1.4.2 The Seller's Optimization Problem

After solving the Buyer's problem, I formalize the Seller's problem. I calculate first the Seller's expected gross profit from choosing an information structure. For any v, the Seller's utility is $U_S(p,v) > 0$ if he offers price p and the Buyer accepts the offer, and $U_S(p,v) = 0$ if the Buyer rejects the offer. As the first mover, the Seller can take into account the Buyer's anticipated response to the price offer. This response is described by the Buyer's conditional probability of acceptance in Equation 1.13. Corollary 1 establishes the gross expected profit of the Seller derived from the Buyer's solution in Proposition 1.

Corollary 1. The Seller's gross expected profit is

$$\int_v \int_p U_S(p,v) f(p,v) \frac{1}{1+exp\left(-\frac{1}{\lambda_B} U_B(p,v)\right)} dpdv \tag{1.14}$$

In summary, the Seller's joint posterior belief would be $f_S(p,v)$ after observing his own signal p, but he can also obtain information indirectly from the Buyer's decision. In case the offer is accepted, the Seller's posterior belief becomes $f_S(p,v)\frac{1}{1+exp\left(-\frac{1}{\lambda_B}U_B(p,v)\right)}$, and if the offer is rejected it becomes $f_S(p,v)\frac{1}{1+exp\left(\frac{1}{\lambda_B}U_B(p,v)\right)}$. From Example 1 one can conclude that the weighting term makes the joint posterior beliefs of the Seller more precise. If the Buyer's signal is to *Accept* the offer, than the probability weights on higher values increase, and on lower values decrease compared to the posterior beliefs of the Seller before making the offer. Similarly if the Buyer's signal is to *Reject* then the probability weights increase for lower values and decrease for higher values.

Alternatively, one can think of the additional weighting term as part of the Seller's utility function, which in this case becomes $U_S(p,v)\frac{1}{1+exp\left(-\frac{1}{\lambda_B}U_B(p,v)\right)}$. Intuitively it means that the Seller's utility is expressed as his expected gain, which depends on the probability of acceptance by the Buyer, as well as on price p and value v. At the same time, the probability of acceptance by the Buyer is also a function of price p and value v. Therefore, the utility function of the Seller maintains its desirable property, that it only depends on p and v, besides the unit cost parameter of the Buyer λ_B. This alternative interpretation also becomes practical while writing the Seller's problem, because the Seller can only choose his own joint posterior distribution $f_S(p,v)$, and pays the cost only of this chosen information. The Seller's information choice problem is summarized as:

Problem (The Seller's Information Choice)**.**

$$\max_{f_S} EU_S - C_S \tag{1.15}$$

$$s.t. \quad g(v) = \int_p f_S(p,v)dp \quad \forall v \tag{1.16}$$

$$f_S(p,v) \geq 0 \quad \forall p,v \tag{1.17}$$

where

$$EU_S = \int_v \int_p U_S(p,v) f_S(p,v) \frac{1}{1 + exp\left(-\frac{1}{\lambda_B} U_B(p,v)\right)} dpdv \tag{1.18}$$

is the expected utility of the Seller, and

$$C_S = \lambda_S * I_S(g, f_S) \tag{1.19}$$

is the cost of information for the Seller.

Equation (1.15) is the objective of the Seller. The first term is the gross expected utility of the Seller after choosing the information structure, but before getting any signal, and the second term is the cost of information. The unit cost of information is denoted by λ_S. Equation (1.16) is the condition which ensures that the chosen information structure is consistent with the prior information.

It was shown in preceding literature by Jung et al. [2019], that the optimal joint posterior distribution from the Seller's problem takes non-zero values only for a discrete set of possible prices. It was also shown that when the joint posterior distribution is non-negative for a given price p and value v, then it will be non-negative for any v and the given price p. These latter result suggests that seeking signals that partitions the support of v is never optimal. Proposition 2 provides the first order condition for the joint posterior distribution $f_S(p,v)$ when it is non-zero.

Proposition 2. The solution to the Seller's problem is

$$f_S(p,v) = \frac{\bar{f}_S(p) exp\left(\frac{1}{\lambda_S} U_S^E(p,v)\right) g(v)}{\sum_p \bar{f}_S(p) exp\left(\frac{1}{\lambda_S} U_S^E(p,v)\right)} \tag{1.20}$$

for any $p \in \mathcal{P}$, and $f_S(p,v) = 0$ otherwise, where $U_S^E(p,v) = \frac{U_S(p,v)}{1+exp\left(-\frac{1}{\lambda_B} U_B(p,v)\right)}$ is the Seller's utility adjusted by the anticipated probability of acceptance by the Buyer. The parameters λ_S and λ_B are the unit information costs for the Seller and the Buyer, respectively, $g(v)$ is the prior distribution of the state variable v, and $\bar{f}_S(p) = \int_{v'} f_S(p,v') dv'$ is the unconditional probability that price p is chosen by te Seller.

For the final solution, the unconditional probabilities $\bar{f}_S(p)$ of choosing each price p need to be found as well. Solving this problem analytically is challenging, because the set of prices $\mathcal{P}$ where the Seller's joint posterior distribution is non-zero has to be specified as well. Instead, I present a numerical solution in the next subsection to find the exact solution to the Seller's problem.

1.4.3 Numerical Solution to the Seller's Problem

I utilize the Blahut-Arimoto algorithm [Arimoto, 1972, Blahut, 1972] to determine the explicit solution to the Seller's problem in Proposition 2. The Blahut-Arimoto algorithm finds the optimal unconditional probability $\bar{f}_S(p)$ as the fixed point of a formula derived from Equation (1.20) above by integrating both sides with respect to v:

$$\bar{f}_S(p) = \bar{f}_S(p) \int_v \frac{exp\left(\frac{1}{\lambda_S} U_S^E(p,v)\right) g(v)}{\sum_p \bar{f}_S(p) exp\left(\frac{1}{\lambda_S} U_S^E(p,v)\right)} dv \tag{1.21}$$

where $U_S^E(p,v) = \frac{p-v}{1+exp\left(-\frac{1}{\lambda_B} U_B(p,v)\right)}$ and $\bar{f}_S(p) = \int_{v'} f_S(p,v') dv'$. The only unknown variable in Equation (1.21) is the Seller's unconditional posterior distribution $\bar{f}_S(p)$. Therefore, it is fully defined as its fixed point. The Seller's joint posterior distribution is also fully defined when $\bar{f}_S(p)$ is plugged back into Equation (1.20). Theorem 1 establishes these results.

Theorem 1. The solution the Seller's problem is

$$f_S(p,v) = \frac{\bar{f}_S(p) exp\left(\frac{1}{\lambda_S} U_S^E(p,v)\right) g(v)}{\sum_p \bar{f}_S(p) exp\left(\frac{1}{\lambda_S} U_S^E(p,v)\right)} \tag{1.22}$$

where $\bar{f}_S(p)$ is defined as the fixed point of

$$\bar{f}_S(p) = \bar{f}_S(p) \int_v \frac{exp\left(\frac{1}{\lambda_S} U_S^E(p,v)\right) g(v)}{\sum_p \bar{f}_S(p) exp\left(\frac{1}{\lambda_S} U_S^E(p,v)\right)} dv \tag{1.23}$$

for given support $p \in \mathcal{P}$.

Csiszar [1974] proved that the Blahut-Arimoto algorithm finds the unique fixed point of Equation (1.23). Here, I present the solution in a numerical example. I iterate Equation (1.23) starting form a discrete, but highly refined support of prices $\mathcal{P} = \{p_1, \dots, p_n\}$ where $p_j = p_{max} \frac{i-0.5}{n}$, and uniform probability distribution $P(p_1^{(0)}) = \bar{f}_j^{(0)} = \frac{1}{n}$ for $j = 1, \dots, n$. The continuous prior distribution $g(v) \sim U(0,1)$ I also replace by a discrete uniform distribution, which has support $\{v_i\}$, $i = 1, \dots, n_v$, and probability distribution $P(v_i) = g(v_i) = \frac{1}{n_v}$ for all $i = 1, \dots, n_v$. More details about the Blahut-Arimoto algorithm can be found in Appendix A.3. I illustrate the numerical solution to the Seller's problem on the following example:

Example 2. Assume that the prior distribution of the value of a good is $v \sim U(0,1)$. The Seller and the Buyer interacts according to the game described in Section 1.3.3. The Seller's unit information cost is $\lambda_S = 9$ and the Buyer's unit information cost is $\lambda_B = 9$. The utility of the Seller is $U_S(p,v) = (p - 100 * v)$ and the utility to the Buyer is $U_B(p,v) = (100 * 3/2v - p)$.

In Example 2, I scale the prices and costs by 100 in sake of better readability of the results, otherwise the setup is equivalent to the Market for Lemons with Endogenous Information problem. I find the numerical solution with the Blahut-Arimoto algorithm, using refinement $n = 300$ for the support of the price, and refinement $n_v = 100$ for the support of the value distribution. Since the minimum price is 0 and the maximum price that the Seller would possibly offer is 150, the points in the support of the initial price distribution are $p_j = \frac{j-0.5}{2}$ for $j = 1, \ldots, 300$. Similarly, the value v takes minimum of 0 and maximum of 1, therefore the points in the support of the value distribution are $v_i = \frac{i-0.5}{100}$ for $i = 1, \ldots, 100$. The prior distribution of the value is the discrete uniform distribution $g(v_i) = \frac{1}{100}$ for every i, and the initial unconditional distribution of the prices is also a discrete uniform distribution $\bar{f_j}^{(0)} = \frac{1}{300}$.

As the result of the Blahut-Arimoto algorithm, from the continuous support of the price distribution three price points arise with non-zero probabilities. I denote them by $p_L = 83.75$, $p_M = 100.25$ and $p_H = 121.75$. The probability distribution over the three price points is $\bar{f_S}(p_L) = 0.24$, $\bar{f_S}(p_M) = 0.25$, and $\bar{f_S}(p_H) = 0.51$. These are the unconditional probabilities that each price is offered. For the full solution to the Seller's problem, which is the joint posterior probability distribution $f_S(p, v)$, the unconditional probabilities are plugged back in to Equation (1.22). The support of $f_S(p, v)$ is $\{p_L, p_M, p_H\} \times \{v_1, \ldots, v_{n_v}\}$

Figure 2.3 displays the solution to the Seller's problem with the parameters defined in Example 2. The prior distribution of the Seller is a uniform distribution on $[0, 1]$, its density function is the magenta-colored horizontal line. After observing his signal, the Seller modifies his belief according to the joint posterior distribution $f_S(p, v)$, where p can take three different values, p_L, p_M or p_H. The solid curves on the graph represent the parts of the posterior joint distribution, the blue line is the part after observing p_L, the green line is the part after observing p_M, and the red line is the part after observing p_L. Note that the parts add up to the prior distribution, such that for each v: $f_S(p_L, v) + f_S(p_M, v) + f_S(p_H, v) = g(v)$, therefore these functions are the parts of the same joint posterior distribution. The conditional posterior distributions of the Seller after observing each price one can get after appropriate normalization.

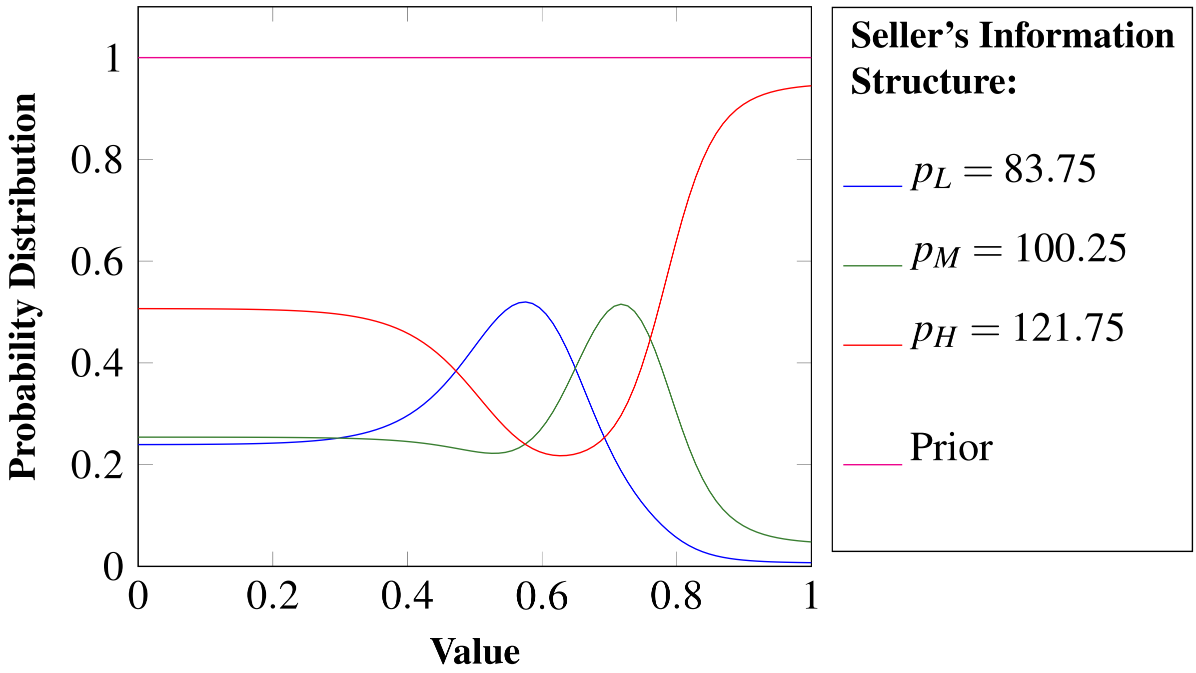


Figure 1.3: The Seller's Optimal Information Structure ($\lambda_S = 9$, $\lambda_B = 9$)

Furthermore, once the result to the Seller's problem is found, the solution to the Buyer's problem is explicitly determined as well for given price offers. I illustrate the solution to the Buyer's problem on Figure 1.4. After observing each price the Buyer gets another signal, either $Accept$ or $Reject$, which is drown from her chosen information structure. The solid curves on each graphs represent the Seller's information choice. The dashed curves are the acceptance probabilities of the Buyer as a function of the value. Note that these functions are not scaled up to be probability distributions. Instead, they represent parts of the joint distribution over three variables, the signal (action) p of the Seller, the signal (action) $Accept$ of the Buyer and the state variable v. This illustration allows us to see the contribution of each price offer to the size of the active market as the shaded areas below the acceptance probability curves of the Buyer.

In Figure 1.5, I illustrate the Seller's conditional posterior distributions and the Buyer's conditional posterior distributions after each observing their own signals. Panel (a) is the Seller's

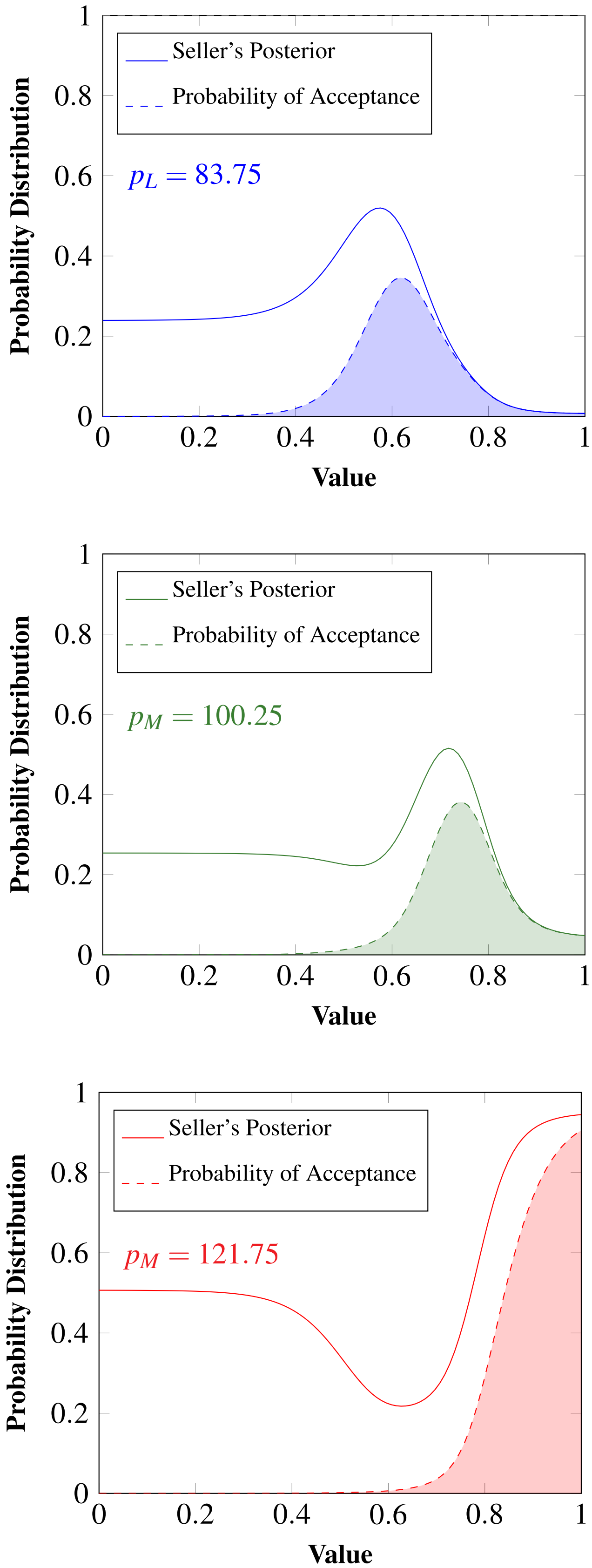
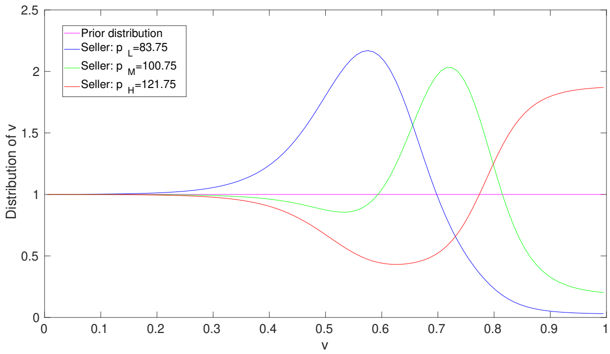
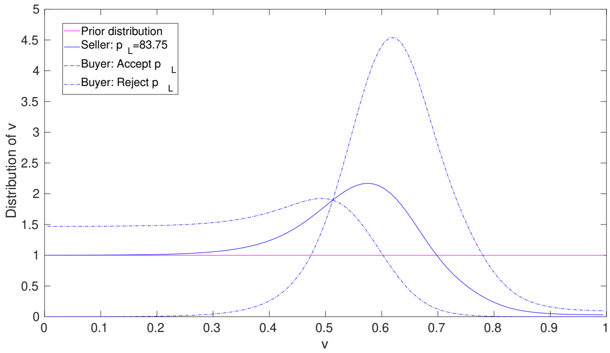


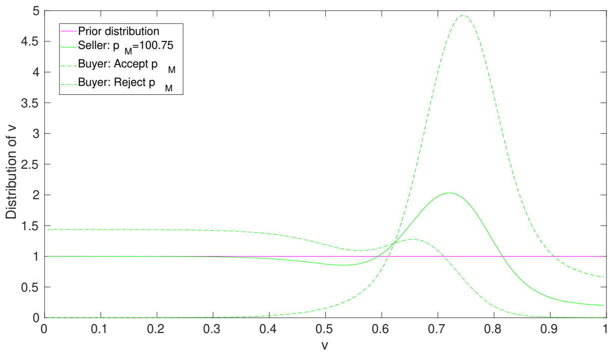
Figure 1.4: The Seller's Price Offer and the Buyer's Response ($\lambda_S = 9$, $\lambda_B = 9$)



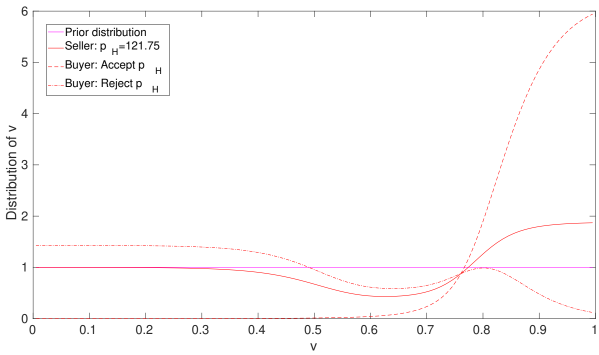
(a) Seller's Posterior Beliefs



(b) Buyer's Posterior Beliefs
$p_L = 83.75$



(c) Buyer's Posterior Beliefs
$p_M = 100.75$



(d) Buyer's Posterior Beliefs
$p_H = 121.75$

Figure 1.5: The Seller's and Buyer's Posterior Beliefs Conditional on Their Signals

conditional posterior distribution, which suggests that the Seller gathers less information about lower values. Close to $v = 0$ the posterior distribution after observing either of the three prices is almost the same as the uniform prior distributions. On the other hand, the Seller gathers more detailed information when v is in the mid-to-high range. However, specifically for the higher values of v, only the highest price signal is distinguished, and the conditional posterior distributions after the two lower price signals are more similar. Panel (b), (c) and (d) of Figure 1.5 illustrates the Buyer's prior distribution, as well as her conditional posterior distributions after observing each price.

1.4.4 Subgame Perfect Equilibrium for Interior Solutions

I characterize the solution described in Section 1.4.1 and Section 1.4.2 as a perfect Bayesian Nash equilibrium. I prove that the interior solution found above indeed forms an equilibrium. I formalize the equilibrium, and provide this result in Theorem 2.

Theorem 2. The following actions and beliefs of the Seller and the Buyer constitute a perfect Bayesian Nash equilibrium of the Market for Lemons with Endogenous Information problem:

1. The Seller chooses information structure $f_S(p,v)$ as described in Section 1.4.2
2. Nature draws the Seller's signal p, the Seller updates his beliefs to $q_S(v|p)$
3. The Seller chooses price p, according to Nature's signal
4. The Buyer observes p, and updates her belief to $q_B^p(v) = q_S(v|p)$
5. The Buyer chooses information structure $f_B^p(D,v)$ according to Section 1.4.1
6. Nature draws Buyer's signal D, the Buyer updates her beliefs to $q_B^p(v|D)$
7. The Buyer observes signal A or R, and accepts or rejects the offer, respectively

To prove that this is a perfect Bayesian Nash equilibrium, one has to show that the Seller's and Buyer's actions are optimal, and that the beliefs are consistent. Let's start with 3. an 7., the actions of the Seller and Buyer after observing their signals. The rational inattention framework ensures that the decision makers always follow their signals, because each signal should be matched with an action. Choosing more signals than the number of possible actions cannot be optimal, because it does not increase the gross utility, but it does increase the information cost. Continuing with the information choices of the Seller and the Buyer 1. and 5., these are results from solving the optimization problem of the Seller and the Buyer. Therefore, these are by definition the best choices in the given environment. The belief updates in 2. and 6. are consistent, which is ensured by the constraints in the Seller's and Buyer's problem. The Buyer's belief update in 4. is also consistent, because the equilibrium specified above is a separating equilibrium. The Seller chooses different prices after different states of the world, which is here defined by the signal sent by Nature in 2.

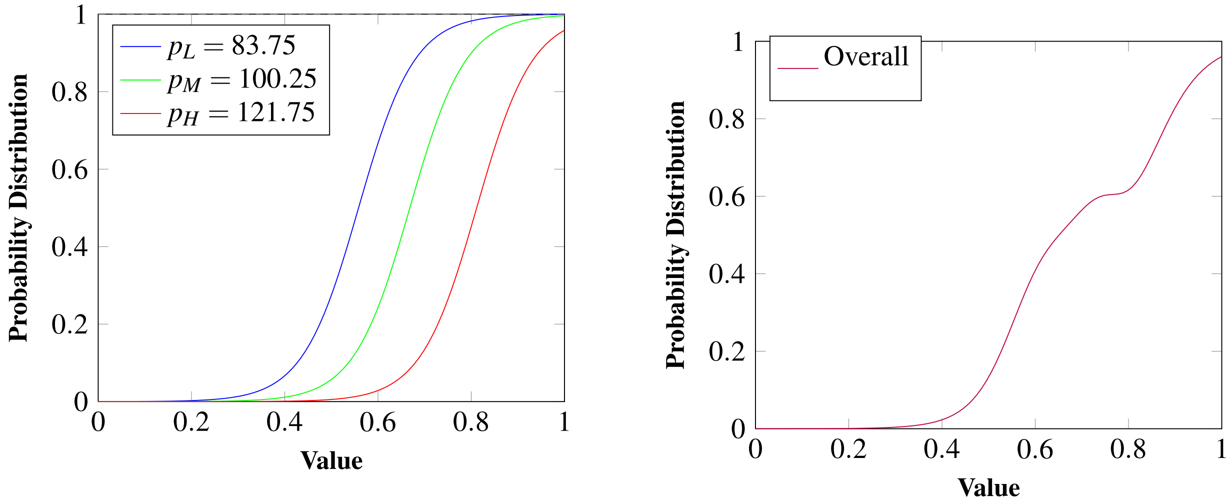
1.4.5 Market Implications

Finally, I summarize some further results from Example 2., and discuss their implications on the market outcomes. First, I calculate the profits for each party. The gross expected utility for the Seller is $U_S = 8.4$ and his information cost is $C_S = 1.13$. Therefore the Seller's net utility is $NetU_S = U_S - C_s = 7.28$. The Buyer's expected utility after observing p_L, p_M or p_H is $U_{BL} = 3.27$, $U_{BM} = 3.76$, or $U_{BH} = 3.92$, respectively. The corresponding information costs are: $C_{BL} = 2.83$, $C_{BM} = 3.19$, $C_{BH} = 3.5$. The gain from gathering information exceeds the cost in all three cases, therefore the Buyer chooses to gather further information, and accepts these prices with positive probabilities. The average expected utility of the Buyer is $U_B = 3.72$ and the average information cost is $C_B^{av} = 3.26$, which results in average net utility $NetU_B = 0.46$

for the Buyer. Taking into account the Seller's and Buyer's utility with equal weights, the total welfare is $NetU = NetU_B + NetU_S = 7.74$.

Compare this to the Akerlof's Market for Lemons setting where under the extreme information asymmetry assumption the market collapses, hence the profit is 0 for both agents. In contrast, my model predicts higher total welfare and positive net profit for both the Seller and the Buyer. However, this welfare is lower than the maximum total welfare 25 that is achieved in the symmetric information benchmark case. Under symmetric information, the market operates efficiently, and the Seller fully enjoys the first mover advantage, hence extracts the entire profit. In my model there is less asymmetry in terms of the outcomes. Out of the total net utility 7.74 the Seller gets around 94% and the Buyer gets around 6%. Therefore, the Seller still gets significantly higher fraction due to his first mover advantage, but the option to the Buyer to gather information herself reallocates some of the profits. In summary, the Buyer is better off, the Seller is worse off and the total welfare is lower in the Market for Lemons with Endogenous Information case than in the symmetric information case.

Second, I analyze the probability of successful selling under my model and compare it to the benchmark settings. In the current model every good is sold with positive probability, however no good is sold with probability 1. The probability of selling a good depends on its quality. The probability that a good with certain quality is sold is displayed in Figure 1.6. The probability of selling increases in the quality for each price as it is shown in Panel (a), and in total as it shown in Panel (b). The average acceptance rate is $P_{AL} = 0.44$ when the price is $p_L = 83.75$, $P_{AM} = 0.33$ when the price is $p_M = 100.75$ and $P_{AH} = 0.19$ when the price is $p_H = 121.75$. The overall average acceptance rate across all prices is $P_A = 0.29$. Comparing this to the benchmark cases, this volume of trade is higher than in the case of extreme information asymmetry, where no trade occurs, but lower than in the case of symmetric information, where all goods are sold



(a) Probability of Acceptance of Each Price

(b) Overall Probability of Acceptance

Figure 1.6: Probability of Selling a Good as the Function of Quality

with probability 1. Therefore, under the current model with endogenous information, the market does not collapse, but there is some level of inefficiency, hence the reduced total welfare above. Furthermore, the average product sold is of higher quality, because the probability of acceptance increases in the quality, as it was shown on Figure 1.6.

1.5 The Effects of Changes in the Information Costs

In this section I investigate the effects of changes in the unit information cost parameters on the individual and market outcomes. I illustrate the movements by simulating the model for a set of different unit information cost parameters for the Seller and the Buyer. I use the case $\lambda_S = 6$ and $\lambda = 6$ as the baseline for comparisons. While analyzing the consequences of changes in the unit information cost parameters, I focus on two distinct effects. First, the level of unit information costs affects the total amount of information in the model, regardless whether the Seller or Buyer acquires it, which has an effect on the market outcomes. Second, the extent of asymmetry between the Seller's and Buyer's information affects the outcomes as well, even when the total amount of information is fixed. In order to better understand the pure effects of information asymmetry in a sense of differing unit information costs, I fix the level of overall information in some of the analyses.

1.5.1 General Observations About the Equilibrium

At first glance, the following observations characterize the effect of the unit information cost parameters on the equilibrium. Figure 1.7 displays the possible price points that the Seller chooses with positive probabilities and the probability distribution over these price points. I fix the unit information cost parameter of the Buyer at $\lambda_B = 6$ and vary the unit information cost of the Seller from $\lambda_S = 6$ to $\lambda_S = 3$. As described above, the Seller optimally restricts his attention to a discrete set of prices. The exact number of the prices offered depends on the magnitude of the unit information cost for the Seller. In particular, the less costly the information is for the Seller, the more price options he uses with positive probability. The range of prices gets wider, as the lowest price decreases and the highest price increases. At the same time, the average price decreases as the unit information cost for the Seller decreases. Table A.1 in Appendix A.2 contains the data used for Figure 1.7.

1.5.2 Information Asymmetry

Second, I look explicitly how the amount of information and the amount spent on information changes with the change in unit information cost parameters. Figure 1.8 illustrates the consequences of the changes in the unit information cost parameters for the Seller on overall information. I show the effects on the amount of information the Seller chooses to gather before offering a price, the amount the Seller spends on this information, the average amount of information a Buyer chooses to gather before accepting or rejecting an offer, and the amount the

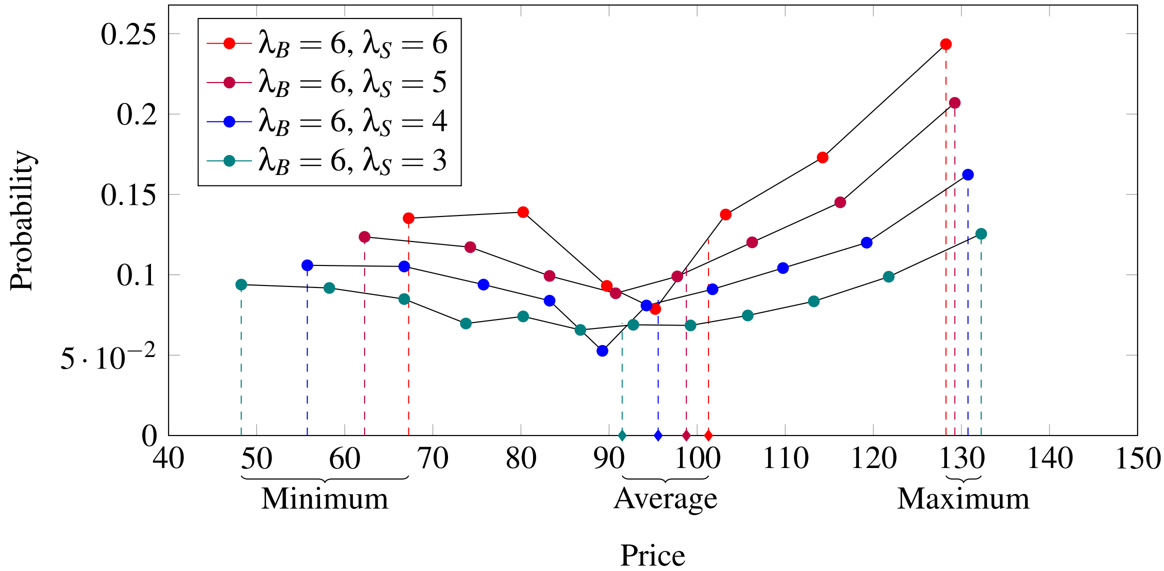
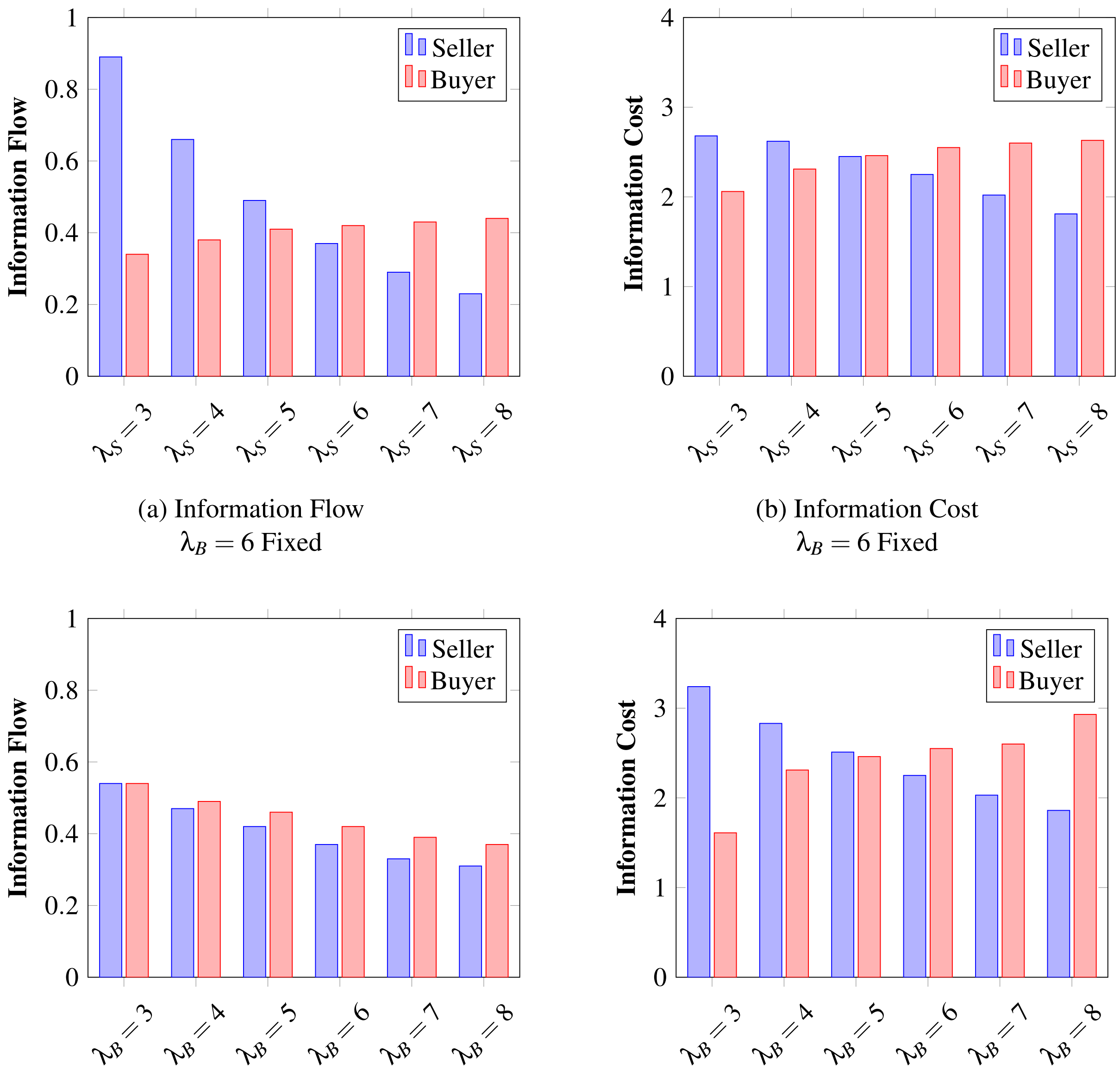


Figure 1.7: Distribution over Possible Price Points

Buyer spends on this information on average. On Panel (a), when the Buyer's information cost is fixed, the Seller's information flow decreases as his unit information cost increases. Asymmetry between the information of the Buyer and the Seller is observable when the Seller's unit information cost is lower than the Buyer's. On Panel (b), it is shown that the Seller also spends less on information when his unit cost of information increases. Both the Buyer's information flow and information cost increases slightly as the unit information cost of the Seller increases.

On Panel (c), when the Seller's information cost is fixed, the information flow for both the Seller and the Buyer decreases as the unit information cost of the Seller increases. Information asymmetry does not occur, because in the cases when the Buyer's unit information cost is higher than the Seller's, the difference is not substantial enough, and it's effect is overweighted by the effect of the Buyer's having more incentives for gathering information. Panel (d) shows that when the Seller's unit information cost is fixed, the Buyer spends more on information as her unit information cost increases. Both of these graphs suggests that the Buyer's demand for information is less elastic than the Seller's. Table A.3 and A.4 in Appendix A.2 contains the data used for Figure 1.8.



(a) Information Flow
$\lambda_B = 6$ Fixed

(b) Information Cost
$\lambda_B = 6$ Fixed

(c) Information Flow
$\lambda_S = 6$ Fixed

(d) Information Cost
$\lambda_S = 6$ Fixed

Figure 1.8: Information Asymmetry Under Varying Unit Information Cost Parameters

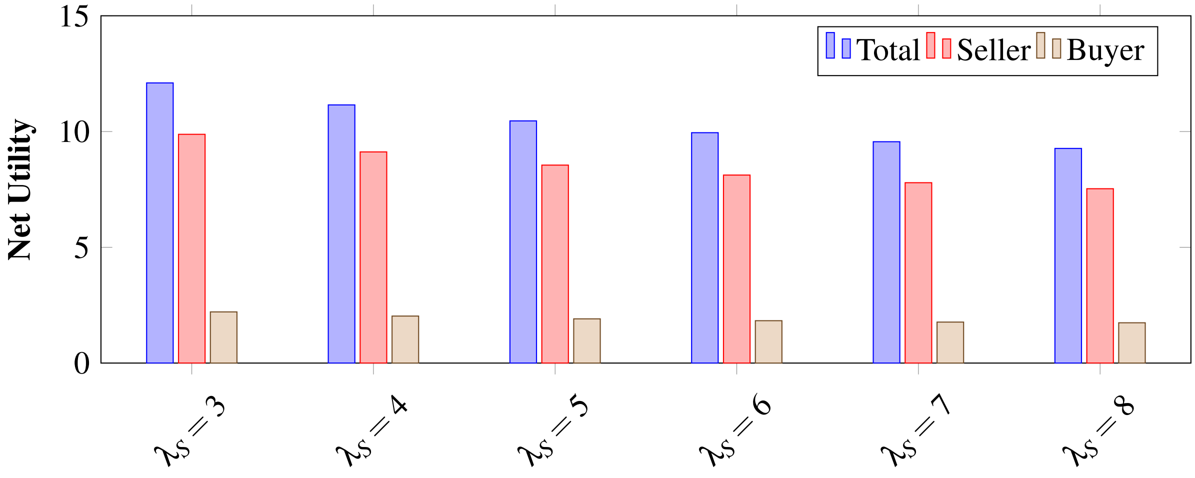
1.5.3 Welfare Consequences

The unit information costs can affect the size and allocation of welfare in two ways. First, the magnitudes of the unit information costs have an effect on the outcomes. Second, the level of asymmetry in the unit information cost parameters of the Seller and the Buyer also has an effect on the outcomes. In Figure 1.9, the gray columns are the changes in the net utility of the Buyer, the pink columns are the changes in the net utility of the Seller. The blue columns are the changes in the sum of the Seller's and Buyer's net utilities, which is the total welfare in this case. Panel (a) shows the case when the Buyer's unit information cost is fixed, Panel (b) shows when the Seller's unit information cost is fixed, and Panel (c) is the case when the total amount of information gather by either the Seller and the Buyer is fixed.

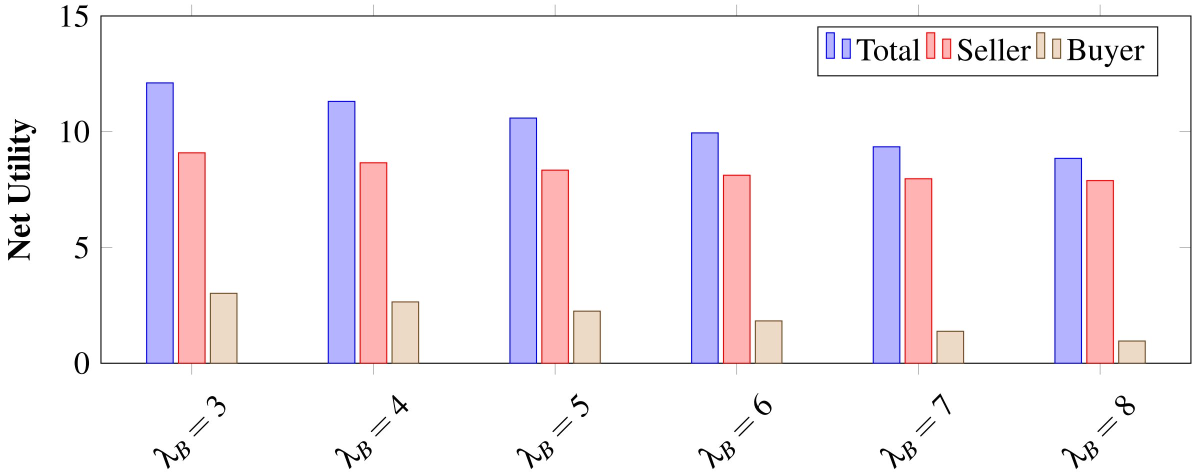
There is an information asymmetry advancing the Seller, when the Seller's unit information cost is lower than the Buyer's unit information cost. In this case, increasing information asymmetry decreases the total welfare, and this decrease is mostly burdens the Buyer. In fact, the Seller is slightly better off due to the information asymmetry in terms of net profit. When the asymmetry advances the Buyer, then total welfare slightly increases, and this change is also mostly on the Buyer's side.

1.5.4 Market Efficiency

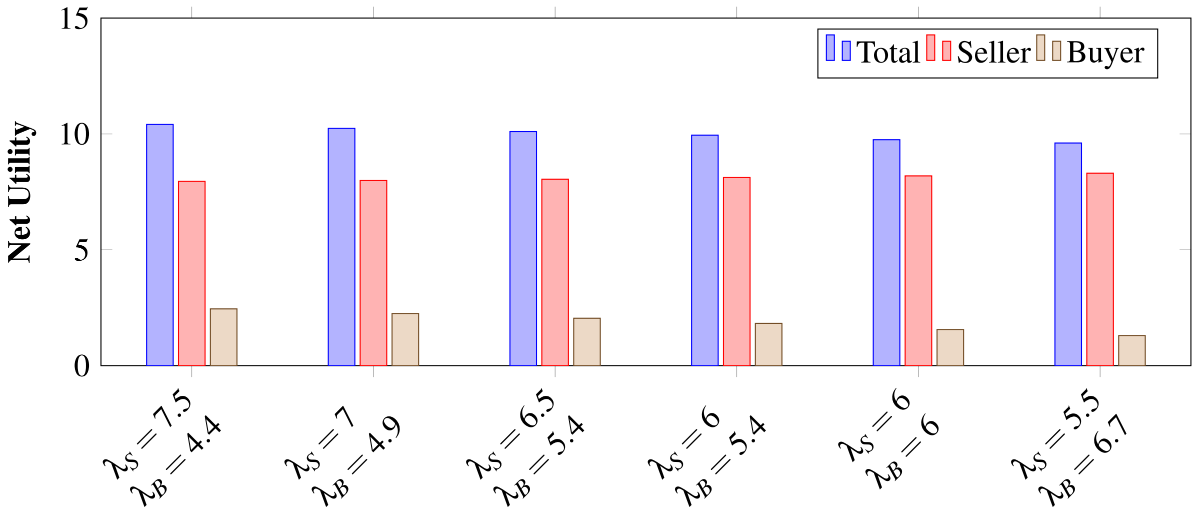
Finally, I show the effect of information costs and information asymmetry on the market efficiency. In the model of Akerlof [1970] the market collapses and no trade occurs. In contrast, in the current model the market still operates, but not at the efficient level. It means that all goods are sold with positive probability, but no good is sold with probability 1. The overall probability that a good is sold is between 0 and 1. In general, it is true that goods of higher quality are more likely to be sold. The policy consequence of this result is that the market efficiency can be



(a) Fixed $\lambda_B = 6$



(b) Fixed $\lambda_S = 6$



(c) Fixed Total Information

Figure 1.9: Change in Net Utilities

increased by decreasing the level of information asymmetry.

1.6 Conclusions

In this paper I build a model of adverse selection under asymmetric information arising from endogenous information gathering. I assume that a Seller and a Buyer, who negotiate over the price of a good of unknown quality, both have access to costly information. The Seller, before offering a price, acquires a signal about the optimal price and the Buyer, after observing the price offer, also acquires a signal whether or not to accept it. The cost of the signals are linear in their ability to reduce the uncertainty of the prior beliefs. The unit costs of information may differ between the Seller and the Buyer, which represents the information asymmetry in this model.

I predict outcomes and market implications within the range of parameters where the interior solution exists, and both the Seller and the Buyer choose to gather information before taking an action. The current assumptions of information asymmetry with endogenous information results in four main finding. First, I conclude that the market does not collapse, however operates with some inefficiencies. Any quality of good sells with probability, but no good sells with certainty. The probability of selling a good depends on the unit information costs, and in general, it is increasing with the quality of the good. The latter finding also implies that the average quality of the goods that are sold are higher then the overall average quality.

The total welfare, defined here as the sum of the Seller's and Buyer's net profits, is positive, but below its possible maximum. Moreover, the profit is nonzero for both parties. This is in contrast with the symmetric case, where the Seller extracts all profits, an the Buyer gets nothing in expected value.

Information asymmetry is still observable in the current model. The information setting I use is suitable to characterize the amount of information each agent uses through the uncertainty reduction between their prior and posterior beliefs. I conclude that when the unit information cost is substantially lower for the Seller than the Buyer, then the Seller indeed gathers more information.

I analyze the pure welfare consequences of information asymmetry by fixing the overall information of the Seller and the Buyer, while increasing the gap between their unit information costs. I find that the increase in information asymmetry, even when the total amount of information is fixed, decreases the total welfare. In particular, the Buyer bears most of the welfare losses. Suitable innervation, such as disclosure requirements or centralized databases, are able reduce these welfare losses, and the current project can be the basis to design such policies.

Chapter 2: Overconfidence, costly information and costly information distortion

2.1 Introduction

Overconfidence – the overestimation of one's own skills, placement or information – has been widely investigated in economics for its impact on the outcomes of individual decision making. An overconfident CEO may over-invest because she overestimates the returns of the investment [Malmendier and Tate, 2015]. A principal who overestimates her own qualities may undervalue an agent, and hence offer sub-optimal contract for both [Heidhues et al., 2018]. Besides economics, overconfidence has an impact on several other aspects of everyday life, such as health, education and workplace [Dunning et al., 2004]. Moreover, the individual biases show their effects on the aggregate level as well. The high rate of business failure is explained by the excess entry of firms who are overconfident about their own potential compared to others [Camerer and Lovallo, 1999]. Overconfident consumers underestimate their future consumption, and certain firms who offer payment plans, such as mobile phone or car rental companies, tend to exploit this naivete in their multi-part pricing schemes [Grubb, 2015].

Overly confident thoughts and behaviors are quite common among the population [Harris and Hahn, 2010], however not all people are equally affected. In particular, the relationship between individuals' skill levels and their overconfidence has been noted for a long time. As Russell [1933] gets straight to the point, "the fundamental cause of the trouble is that in the modern world the stupid are cocksure while the intelligent are full of doubt." In a more scientific setting, Kruger and Dunning [2000] provided the first empirical evidence that those who are less skilled are indeed the ones who are less likely to be aware of it. The unskilled-and-unaware effect – also known as the "Dunning-Kruger effect" after its first investigators – refers to the inverse relationship between one's own actual and perceived performance.

The Dunning-Kruger effect has not been frequently discussed in economic settings, although it is closely linked to information, learning, and their relationship to overconfidence. In a recent study Sanchez and Dunning [2018] tracked the actual and perceived performance of subjects who were initially new to a certain task, but who were given the opportunity to learn from repeated trials and feedback. The task was designed to take place in a fictional environment, such that subjects could not use prior expertise. Results indicated that after an initial jump – what the authors called the "beginners' bubble" – overconfidence gradually declined as the subjects gained more experience. This result suggests that overconfidence occurs not as a fixed personal attribute, but as a feature that can be changed over time. Hence, learning can mitigate both being unskilled and being unaware of it. Consequently, the Dunning-Kruger effect is related to – and should be analyzed together with – learning and the change in skill level.

A study by Sawler [2021] noticed the relevance of the Dunning-Kruger effect in education and proposed potential solutions to reduce its disadvantageous consequences. In his experiment he showed that those who are taking only introductory level economic courses are likely to remain overconfident about their knowledge of the field. The psychology literature thoroughly discusses this connection in further examples [Dunning, 2011], however none of the studies before provide a mechanism through which the improvement of skills may affect overconfidence. In this paper, I propose a model that explains the Dunning-Kruger effect as being a result of a constrained utility maximization problem, where the decision maker is rationally inattentive to information, and at the same time biased toward more optimistic subjective beliefs. The setup is a tractable form to augment into various economic applications, where the formation of overly confident subjective beliefs different from the objective information may have an effect on the final outcomes.

My model builds on two distinct approaches from economic theory, and extends these methodologies by one additional assumption. First, I use a framework where decision makers are allowed to choose their objective information and subjective information separately, an idea that appears in Brunnermeier and Parker [2005]. Furthermore, following the literature on motivated beliefs by Bénabou and Tirole [2016], I assume that overly optimistic beliefs have their intrinsic values, therefore the individuals may choose to be overconfident even when it comes with a loss in materialistic terms.

Second, I take into account two types of costs for choosing objective and subjective information. I measure the cost of the objective information by entropy-based cost functions grounded in the rational inattention literature (summarized by Mackowiak et al. [2020]). Similar to the experiment of Sanchez and Dunning [2018], my model describes repetitions of an identical task. In each task, the decision maker can gather an arbitrary amount of information before taking an action. In accordance with the assumptions of the rational inattention framework, choosing more information provides higher expected utility, but also imposes a cost. I assume that this cost is decreasing with the skill level of the decision maker, which is increasing through the learning process.

Finally, as a third assumption, I introduce the cost of information distortion. Brunnermeier and Parker [2005] suggested that the only constraint of choosing overly optimistic beliefs is the consequent future monetary loss. In contrast, I assume that forming an overconfident subjective belief imposes direct cognitive costs. In line with the previous parts of the model, I define the cost of information distortion within the entropy-based rational inattention framework, but unlike the cost of objective information, I assume that it increases with the skill level. That means the more skilled the decision maker is, the easier she solves an upcoming task, and at the same time, the harder she misjudges her own performance.

These assumptions are aligned with the explanation by Dunning [2011] for the existence of the unskilled-and-unaware phenomenon, however my interpretation is slightly different. In his argument, he connected ignorance with meta-ignorance, and showed that the exact same skills are necessary to perform well as to judge accurately one's own performance. He calls overconfidence a result of a double burden of incompetence; to be unskilled, and unaware about being unskilled at the same time. On the contrary, my model takes a rational approach. I assume that overconfidence is the result of maximizing a weighted sum of a materialistic expected utility component and a belief utility component, net of the cost of information and the cost of information distortion.

My project contributes to the body of behavioral economic models that connect information processing and overconfidence. Mobius et al. [2011] showed that the misinterpretation of signals leads to overconfident beliefs. Eil and Rao [2011] argue that overconfidence may prevent the processing of objective information, when its content is unpleasant. The current project also proposes a model that connects overconfidence and learning, but with two distinctions. First, in my setup I assume that the overly optimistic beliefs are the result of the decision maker's utility-maximizing choice. Second, instead of an exogenous source of information, I assume endogenous information acquisition. The choice of using endogenous information also establishes a bridge between overconfidence and the rational inattention framework. My model extends the existing literature on rational inattention by adding a behavioral bias toward more optimistic subjective beliefs to a costly information acquisition problem.

In addition to the contribution to the economic theory literature, the project has potential for policy applications. For example, in Chapter 3, I show as a consequence of my model that the individuals with fewer skills are the ones who are less willing to improve. More specifically, until a certain threshold, increasing skills results in a loss in utility, therefore there is no internal motivation for learning. Hence, the initial "beginners' bubble" of overconfidence found by Sanchez and Dunning Sanchez and Dunning [2018] is not only harmful in current decisions, but also ruins incentives for further improvement. This argument highlights the importance of nudging the earlier stages of learning, for example, through mandatory education or training programs. However, after a certain threshold, no further external incentive is needed, and the learning process continues voluntarily. At that point, the trouble identified by Bertrand Russell [1933] might turn into an advantage; those who are "intelligent but full of doubt" become eager to learn more.

2.2 Literature Review

This project contributes to the literature on overconfidence in economics by offering a model which explains the relationship between skill levels and overconfidence by the simultaneous endogenous choice of costly information and subjective beliefs. The precise terminology of overconfidence was introduced by Moore and Healy [2008], and later revised by Moore and Schatz [2017]. They determined three distinguished categories of overly optimistic beliefs. It is overestimation of one's own performance when someone thinks herself better than she is. Overplacement is when someone exaggeratedly thinks herself better than others. Finally, overprecision occurs when someone thinks that her information is more accurate than it actually is. The current model captures a case of overprecision, because the decision maker's overly optimistic assessment is about her own level of informedness.

Optimism, in a broader sense, was recognized in psychology literature for its potential benefits on well-being starting from Harris and Hahn [2010], and more recently summarized by Sharot [2011]. Overconfidence in economics and finance, however, is of primary interest because it can lead to sub-optimal decisions in the materialistic sense. All three categories of overconfidence appears in applications in finance by Barber and Odean [2001], Bhandari and Deaves [2006], and Gervais et al. [2011], in economics by Camerer and Lovallo [1999], de la Rosa [2011], Heidhues et al. [2018], and Malmendier and Tate [2015] and in political economy by Ortoleva and Snowberg [2015]. These applications identify overconfidence as an existing bias in decision making, but are vague about its sources.

The literature on the economics of motivated beliefs suggests the reasons for, and the mechanism behind overconfident thoughts and actions. These models, following the observations of psychology, assume that overly optimistic beliefs are not just a mistake in rational decision making, but also are desirable outcomes, which are part of the decision maker's objective. Examples of these models are found in Bénabou and Tirole [2002] and Caplin and Leahy [2001]. A recent survey by Bénabou and Tirole [2016] reviews how these motivated beliefs, in general, can be used as a direct or instrumental source of utility.

In summary, overconfidence in economics has been recognized both for its benefits and caveats, and in my model both effects are recognized. Closest to the current paper, Brunnermeier and Parker [2005] proposed a model in which there is a trade-off between materialistic and belief based utilities. In their setting, an investment decision is made under uncertainty, and welfare arises both from the expected utility of the investment in monetary terms and from the belief about this future income. One of their key idea is the separation of subjective beliefs from objective beliefs, which I apply similarly in the current project. However, my setup is different in two ways. First, I replace the exogenous information source to an endogenous costly information acquisition process. Second, I incorporate a direct cognitive cost element for forming overly optimistic beliefs besides its adverse effect on materialistic outcomes.

Beyond the positive and negative consequences of overly optimistic beliefs, the bi-way relationship between overconfidence and information has been also analyzed in economics. On one hand, overconfidence may effect information processing and learning. Heidhues et al. [2018] showed that the biased perception of signals due to overestimation of the agent's own qualities leads to sub-optimal outcomes of a contract. On the other hand, learning might be the way to mitigate behavioral biases. Papers by Mobius et al. [2011] and by Eil and Rao [2011] discussed the evolution and possibilities of decreasing overconfidence by trials and feedback. These models commonly used fixed exogenous sources of signals, therefore they all differ from my setup. Moreover, the current project directly analysis the effect of the information acquisition process on beliefs, rather than the effect of the available information. Therefore, my model is suitable to answer questions specifically about the inference of information gathering and overconfidence.

Parallel to the investigation in economics, researchers in psychology also noted the relationship between learning and overconfidence. The first empirical analysis was provided by Kruger and Dunning [2000]. In their experiment they showed the inverse relationship between one's actual ranking and one's estimated ranking in a competition. The gap between the estimated ranking and the actual ranking is an instance of overplacement, which was empirically proven to be decreasing as the actual ranking increased. Since then, Dunning-Kruger effect is the broad terminology used in psychology for the inverse relationship between skills and overconfidence. A series of additional studies, such as Krueger and Oakes Mueller [2002] and Kruger and Dunning [2002], continued the investigation of the phenomena.

All the above mentioned early studies focused exclusively on between-subject comparisons, but later Sanchez and Dunning [2018] found evidence of the Dunning-Kruger effect in a within-subject design as well. In their experiments they investigated learning with little initial skills, and showed that after gaining a minimal experience overconfidence jumps to its maximum state and then gradually decreases from that pike. I propose a theoretical explanation of these empirical findings, focusing on the decrease of overconfidence with the increase of skill level. My quantitative model is in line with the theoretical argument of Dunning [2011]. In his study he connected ignorance with meta-ignorance, and showed that the exact same skills are necessary to perform well as to judge accurately one's own performance.

The literature on Dunning-Kruger effect also formalizes ideas that have appeared in earlier models describing the process of learning. For example, the popular "four stages of learning" scheme, summarized and connected to the Dunning-Kruger effect by Adams [2017], was originally developed by Noel Burch in the 1970s . In his setup, the individual starts from being "Unconsciously Unskilled", then she turns first into "Consciously Unskilled", then into "Consciously Skilled", and finally into "Unconsciously Skilled". At the first stage, overconfidence is at its maximum, and starts to decrease in the second stage. Then, in the third stage confidence starts matching the actual skills, and in the last stage the individual can deliver high performance even without thinking of it. Therefore, the first two stages of this learning scheme are perfectly in line with the later formalization of the Dunning-Kruger effect.

A recent paper by Sawler [2021] highlights and validates in an experiment further consequences of learning with overconfidence. He observes that many undergraduate students who take introductory economics, but do not continue on more advanced courses, do not grasp the complexity of the field, and hence remain overconfident about their own knowledge. His paper also argues the possible connection between the Dunning-Kruger effect and the managerial overconfidence described by Malmendier and Tate [2015]. The proposed solution to the overconfidence of the undergraduates is to assign readings that foreshadow the issues beyond the course level. In contrast, in the application of my model I target a similar question, however I propose an alternative solution by lengthening the mandatory phase of learning.

Lastly, this project also contributes to the growing body of work on attention and information acquisition. Models and applications of inattention in behavioral economics are summarized in Gabaix [2019]. Golman et al. [2017] surveys models where information is part of the objective function of the decision maker. More specifically, my model utilizes the framework of entropy-based rational inattention first proposed by Sims [2003b], and later characterized by Matejka and McKay [2011], Caplin and Dean Caplin and Dean [2015] andCaplin et al. [2017]. Experimental tests of the theory can be found in Dean and Neligh [2017]. The rational inattention framework suggests that attention is associated with a cognitive cost, hence the decision maker chooses the optimal amount of information based on a constrained optimization problem. Compared to this framework, I use an additional entropy-based function to measure the cost of information distortion which occurs when the decision maker forms false subjective beliefs.

A recent survey by Mackowiak et al. [2020] collects a wide range of projects utilizing the rational inattention framework. Most of these works are either applications with fully rational, albeit inattentive decision makers, or models aiming to explain biased decisions as a consequence of rational inattention. My project is among the first few setups, where the rational inattention framework is combined with a behavioral bias. Previously, Pagel [2018] incorporated loss-aversion, while in my model the decision maker is biased toward overly optimistic subjective beliefs.

2.3 Model

In this section I describe a model that explains the decrease in overconfidence as the skill level increases – the Dunning-Kruger effect – as a result of a constrained utility maximization problem with costly information acquisition. In this problem the decision maker chooses her objective and subjective information separately in order to maximize her expected utility net of the costs of information. The gross utility stems from two different sources, a materialistic utility component and a belief utility component.

The materialistic utility component is the expected material income from a risky investment decision, where the expectation is formed based on the decision maker's objective information. While the belief utility is the expected income from the same risky investment, but based on the, possibly optimistically distorted, subjective information. The overall utility which is taken into account in the objective function is the weighted sum of the two utility components.

The decision maker also faces two types of costs related to her information choice. On one hand, the choice of the level of objective information imposes information cost. Following the entropy-based rational inattention literature (as formalized in Caplin and Dean [2015]), the cost of information is measured by a linear function of the mutual information between the prior belief and the objective posterior belief of the decision maker. The unit information cost parameter is assumed to be a decreasing function of the skill level. On the other hand, I assume that a subjective information that is different from the objective information imposes an additional cost component, which I call the cost of information distortion. I measure the cost of information distortion by a linear function of the mutual information between the objective and subjective beliefs. The unit cost of information distortion is assumed to be an increasing function of the skill level.

In this model overconfidence is represented by the gap between the subjective and objective information. Overly confident beliefs, although they result in lower materialistic utility, are motivated by the belief utility component. The choices of the unit information cost and unit information distortion cost parameters reflect the explanation of the Dunning-Kruger effect by Dunning [2011]. The higher skill level allows the decision maker to gather information at a lower cost, but at the same time, it makes the distortion to subjective information more costly for her. The balance of the gains and losses due to the two utility components and the two types of costs shapes the decreasing relationship between the skill level and overconfidence.

2.3.1 Cost of Information and Cost of Information Distortion

I introduce the model of overconfidence on a simple two-stage decision problem. The decision maker's prior information is $\boldsymbol{\mu} = \binom{\mu}{1-\mu}$, where μ is the probability of State 1 in a binary-state world. In the first stage the decision maker simultaneously chooses objective information $\boldsymbol{p}$ and subjective information $\boldsymbol{q}$. $\boldsymbol{p}$ is an information structure, which consists of two signals. The first signal generates posterior belief $\boldsymbol{p^a} = \binom{p}{1-p}$ over the two possible states of the world, and is sent with probability $Q(\boldsymbol{p^a})$. The second signal generates posterior belief $\boldsymbol{p^b} = \binom{1-p}{p}$, and is sent with probability $1 - Q(\boldsymbol{p^a})$. The probability $Q(\boldsymbol{p^a})$ is such that this information structure is consistent with the prior distribution. Subjective information $\boldsymbol{q}$ is a distorted version of the objective information $\boldsymbol{p}$. Choosing subjective information structure $\boldsymbol{q}$ means that the decision maker's subjective posterior belief is $\boldsymbol{q^a} = \binom{q}{1-q}$ when her objective posterior belief is $\boldsymbol{p^a}$, and $\boldsymbol{q^b} = \binom{1-q}{q}$ when her objective posterior belief is $\boldsymbol{p^b}$. In the second stage of the problem, after observing her signal, the decision maker chooses between two actions. Action a pays R in the first state of the world, and nothing in the second state of the world, while action b pays nothing in the first state of the world and R in the second state of the world.

The first assumption states that the decision maker is free to choose objective and subjective information separately. This assumption enhances the entropy-based rational inattention model with the option of overconfident beliefs. More specifically, overconfidence arises as a gap between subjective and objective information.

Assumption 1. Given prior information $\boldsymbol{\mu}$, the decision maker chooses objective information structure $\boldsymbol{p}$ and subjective information structure $\boldsymbol{q}$ separately. Signal $\boldsymbol{p^a}$ generates objective posterior belief $\binom{p}{1-p}$ and subjective posterior belief $\binom{q}{1-q}$. Signal $\boldsymbol{p^b}$ generates objective posterior belief $\binom{1-p}{p}$ and subjective posterior belief $\binom{1-q}{q}$. Signal $\boldsymbol{p^a}$ is observed by probability $Q(\boldsymbol{p^a})$, and signal $\boldsymbol{p^b}$ is observed by probability $1-Q(\boldsymbol{p^a},)$, where $Q(\boldsymbol{p^a})$ is such that both information structures $\boldsymbol{p}$ and $\boldsymbol{q}$ are consistent with the prior belief $\boldsymbol{\mu}$.

By choosing objective and subjective information the decision maker aims to maximize her objective function, which consists of two utility components and two types of costs. I denote the materialistic utility component by EU_p, which is the expected utility from the investment in the second stage considering the objective information $\boldsymbol{p}$. Similarly, EU_q denotes the belief utility component, and it is the expected utility from the same second stage investment when the decision maker considers subjective information $\boldsymbol{q}$ as her posterior beliefs. The gross expected utility is a weighted sum of the two components, where the weight on the materialistic expected utility is θ, and the weight on the belief utility is $1-\theta$. In terms of costs, I denote the cost of information by $\boldsymbol{C}$, which is a function of the mutual information between the prior information $\boldsymbol{\mu}$ and the objective posterior information $\boldsymbol{p}$. The cost of information distortion is denoted by $\boldsymbol{D}$, and it is a function of the mutual information between the objective posterior information $\boldsymbol{p}$ and the subjective posterior information $\boldsymbol{q}$. The second assumption summarizes the decision maker's objective function.

Assumption 2. For given prior information $\boldsymbol{\mu}$, the decision maker solves

$$\max_{\boldsymbol{p},\boldsymbol{q}} \theta EU_p + (1-\theta)EU_q - \boldsymbol{C} - \boldsymbol{D} \tag{2.1}$$

where EU_p is the materialistic expected utility component, EU_q is the belief utility component, $\boldsymbol{C}$ is the cost of information, $\boldsymbol{D}$ is the cost of information distortion, and $\theta \in [0,1]$ is a weighting parameter. $\boldsymbol{C}$ is a function of the mutual information between prior information $\boldsymbol{\mu}$ and objective posterior information $\boldsymbol{p}$ and $\boldsymbol{D}$ is a function of the mutual information between objective posterior information $\boldsymbol{p}$ and subjective posterior information $\boldsymbol{q}$.

The next assumption specifies constraints of choosing objective information $\boldsymbol{p}$ and subjective information $\boldsymbol{q}$. The constraint on the objective information is the "No Information Disposal" principle. This constraint states that the information content of the objective information cannot be less than the information content of the prior information. The second constraint I call the "No Underconfidence" principle. It states that the information content of the subjective information is not less than the information content of the objective information. More formally, since the information content is measured based on the mutual information between distributions, information structure $\boldsymbol{p^1} = \left(\binom{p^1}{1-p^1}, \binom{1-p^1}{p^1}\right)$ contains more information than $\boldsymbol{p^2} = \left(\binom{p^2}{1-p^2}, \binom{1-p^2}{p^2}\right)$ if $p^1 > p^2$. Furthermore, before the formal statement of the assumption on the constraints, let us introduce a notational convention. Assume that $\mu \geq \frac{1}{2}$, meaning that the notation of the prior belief $\boldsymbol{\mu} = \binom{\mu}{1-\mu}$ is such that the first state of the world is always the one which initially occurs more frequently.

Assumption 3. For any prior information $\boldsymbol{\mu} = \left(\binom{\mu}{1-\mu}, \binom{1-\mu}{\mu}\right)$ where $\mu \geq \frac{1}{2}$, the objective information $\boldsymbol{p} = (\boldsymbol{p^a}, \boldsymbol{p^b}) = \left(\binom{p}{1-p}, \binom{1-p}{p}\right)$ is chosen under the restriction that $p \geq \mu$ ("No Information Disposal”). For any objective information $\boldsymbol{p}$, the subjective information $\boldsymbol{q} = \left(\boldsymbol{q^a}, \boldsymbol{q^b}\right) = \left(\binom{q}{1-q}, \binom{1-q}{q}\right)$ is chosen under the restriction that $q \geq p$ (“No Underconfidence”).

As a consequence of the third assumption and the notational convention $\mu \geq \frac{1}{2}$, in the second stage of the decision problem the decision maker maximizes her expected utility by choosing action a after getting signal $\boldsymbol{p^a}$, and by choosing action b after getting signal $\boldsymbol{p^b}$. This is because for a given signal, regardless the cost of information and information distortion the choice of the objective and subjective information imposed in the first stage, it gives the highest utility to choose a when the first state is more likely to occur, and choose b when the second state is more likely to occur.

Assumptions 4. and 5. describes the functional forms of the information cost and the information distortion cost, and connects them to the individuals’ skill levels. In accordance with the rational inattention framework characterized in Caplin and Dean [2015], I use entropy-based cost functions both for the cost of information and for the cost of information distortion. More specifically, the cost of the objective information is proportionate to the Shannon mutual information between the objective information and the prior information, and the cost of the information distortion is proportionate to the Shannon mutual information between the subjective information and the objective information.

Formally, given prior information $\boldsymbol{\mu}$ and chosen objective information $\boldsymbol{p}$, the cost of information is

$$c(\boldsymbol{\mu}, \boldsymbol{p}) = \lambda(H(\mu) - (Q(\boldsymbol{p^a})H(p) + (1 - Q(\boldsymbol{p^a}))H(1-p))) \tag{2.2}$$

where $H(p) = -p\log p - (1-p)\log(1-p)$ is the Shannon entropy of the posterior distribution $\binom{p}{1-p}$, and λ is the unit cost of information. After simplifications, the cost of objective information $\boldsymbol{p}$, given prior information $\boldsymbol{\mu}$ is

$$c(\boldsymbol{\mu}, \boldsymbol{p}) = \lambda(H(\mu) - H(p)) \tag{2.3}$$

Similarly, given objective information $\boldsymbol{p}$ and subjective information $\boldsymbol{q}$, the cost of information distortion is

$$d(\boldsymbol{p}, \boldsymbol{q}) = \gamma(H(p) - H(q)) \tag{2.4}$$

where γ is the unit cost of information distortion.

Assumption 4. The cost of information is defined as

$$c(\mu, p) = \lambda(\kappa)(H(\mu) - H(p))) \tag{2.5}$$

where $H(p) = -p\log p - (1-p)\log(1-p)$ is the Shannon entropy of the posterior distribution $\binom{p}{1-p}$, κ is a parameter that measures skill level, and $\lambda(\kappa)$ is the unit information cost, which is a decreasing function of κ with $\lambda(\kappa) \to \infty$ as $\kappa \to 0$, and $\lambda(\kappa) \to 0$ as $\kappa \to \infty$

As an example I apply unit information cost $\lambda(\kappa) = \frac{1}{\kappa}$. This functional form is supported by the cumulative average model of learning by Wright [1936]. According to Wright, the learning curve can be described as $Y = aX^b$, where Y is the cumulative average time (or cost) per unit, X is the cumulative number of units produced, a is the time or cost required to produce the first unit, and b is the slope of the function when plotted on log-log paper. For simplicity, let's set $a = 1$, $b = -1$. Let X be the current skill level, further on denoted by κ. Let Y be the (average) cost of information, analogously the unit cost of information, denoted by λ. Therefore, the formula that describes the inverse relationship between the current skill level of the decision maker and the unit cost of information is

$$\lambda(\kappa) = \kappa^{-1}. \tag{2.6}$$

Assumption 5. The cost of information distortion is defined as

$$d(p,q) = \gamma(H(p) - H(q)) = \beta\gamma(\kappa)(H(p) - H(q)) \tag{2.7}$$

where and $H(p) = -p\log p - (1-p)\log(1-p)$ is the Shannon entropy of the posterior distribution $\binom{p}{1-p}$, κ is a parameter that measures skill level, β is the awareness parameter, and (with a slight abuse of notation) $\gamma = \beta\gamma(\kappa)$ is the unit cost of information distortion. $\gamma(\kappa)$ is an increasing function of κ and $\gamma(\kappa) > 0$ for every κ.

In the example I use $\gamma(\kappa) = \kappa$, hence the cost of information distortion is $\beta\gamma(\kappa)$. A notable difference between the cost of information and the cost of information distortion is that the former decreases, but the latter increases in the skill level. This assumption is crucial in the explanation of the Dunning-Kruger effect, and it is in line with the theory of Dunning [2011], that the more unskilled an individual is, the more unaware she is about it. Or, using the terminology of the current model, the lower skill level enables more inaccuracy in one's own self assessment. As the skill level increases, the increasing costs of information distortion makes it harder to misjudge one's own performance, and forces the decision maker to form more realistic subjective beliefs.

The parameters κ, β and functions $\lambda(\kappa)$ and $\gamma(\kappa)$ defined above describe the decision maker's individual abilities and preferences in a certain decision problem. Although these parameters are all fixed in a specific decision problem, some of them can change over time, across repeated tasks. In particular, the increase of κ represents the learning process, which results in the increase of skill level. The awareness parameter β represents the personal "taste" for distorting information, and it is fixed in this model. Moreover, the cost of information λ and the cost of information distortion γ are function of κ, therefore they also change with κ through the learning process. The choice variables in the decision maker's problem we are focusing on are the objective information p and the subjective information q. A decision problem is defined for given κ, therefore p and q also change with the skill level.

Given the current level of skills (κ) and awareness parameter (β), the following model formalizes the utility maximization problem of the decision maker.

Model (Overconfidence)**.** The decision maker solves

$$\max_{p,q} \theta EU_p + (1-\theta)EU_q - c(\mu, p) - d(p,q) \tag{2.8}$$

$$s.t. \quad p \geq \mu \tag{2.9}$$

$$and \quad q \geq p \tag{2.10}$$

where

$$c(\mu, p) = \lambda(\kappa)(H(\mu) - H(p)) \tag{2.11}$$

$$d(p,q) = \beta\gamma(\kappa)(H(p) - H(q)). \tag{2.12}$$

I assume risk neutral decision maker with linear utility function $u(x) = x$. That means for given objective information $\boldsymbol{p}$ the optimal choice of action in the second stage results in materialistic expected utility $EU_p = pR$, because the decision maker chooses the state which pays R with probability p, regardless which state is it. Similarly, for given subjective information $\boldsymbol{q}$ the optimal choice of action in the second stage results in belief utility $EU_q = qR$. I also assume $\mu = \frac{1}{2}$, which reflects a case where the decision maker has no information about the state of the world before collecting information, hence her prior distribution is discrete uniform. Under these assumptions, the solution to the Overconfidence Model is

$$p = \begin{cases} \frac{1}{1+2^{-\frac{\theta R}{\lambda(\kappa)-\beta\gamma(\kappa)}}} = p^*, & \text{if } \frac{(1-\theta)\lambda(\kappa)}{\beta\gamma(\kappa)} \geq 1 \\ \frac{1}{1+2^{-\frac{R}{\lambda(\kappa)}}} = \overline{p}, & \text{if } \frac{(1-\theta)\lambda(\kappa)}{\beta\gamma(\kappa)} < 1 \end{cases} \tag{2.13}$$

$$q = \begin{cases} \frac{1}{1+2^{-\frac{(1-\theta)R}{\beta\gamma(\kappa)}}} = q^*, & \text{if } \frac{(1-\theta)\lambda(\kappa)}{\beta\gamma(\kappa)} \geq 1 \\ \frac{1}{1+2^{-\frac{R}{\lambda(\kappa)}}} = \overline{p}, & \text{if } \frac{(1-\theta)\lambda(\kappa)}{\beta\gamma(\kappa)} < 1 \end{cases} \tag{2.14}$$

Implications of the results are discussed in Section 2.4. For proof see Appendix B.1.

Further on, I will assume $\mu = 1/2$, $\theta = 1/2$ and $R = 1$. Under these assumptions, the solution of the Overconfidence Model with is:

$$p = \begin{cases} \frac{1}{1+2^{-\frac{\kappa}{2(1-\beta\kappa^2)}}} = p^*, & \text{if } \kappa \leq \frac{1}{\sqrt{2\beta}} \\ \frac{1}{1+2^{-\kappa}} = \overline{p}, & \text{if } \kappa > \frac{1}{\sqrt{2\beta}} \end{cases} \tag{2.15}$$

$$q = \begin{cases} \frac{1}{1+2^{-\frac{1}{2\beta\kappa}}} = q^*, & \text{if } \kappa \leq \frac{1}{\sqrt{2\beta}} \\ \frac{1}{1+2^{-\kappa}} = \overline{p}, & \text{if } \kappa > \frac{1}{\sqrt{2\beta}} \end{cases} \tag{2.16}$$

2.3.2 Model with No Information Distortion

For comparison, I solve the costly information acquisition problem without the option of choosing objective information and subjective information separately. In this case the decision maker still faces cost of information but not the cost of information distortion, since different subjective and objective information is not allowed. I call this model the No Information Distortion Model, and I use it as a benchmark for the Overconfidence Model with information distortion. Formally, I add the restriction $q = p$, but otherwise I use the same assumptions as above. When the unit cost of information $\lambda(\kappa) = \frac{1}{\kappa}$, the restricted model becomes:

Model (No Information Distortion). The decision maker solves

$$\max_{p} EU_p - \frac{1}{\kappa}(H(\mu) - H(p)) \tag{2.17}$$

$$s.t. \quad p \geq \mu \tag{2.18}$$

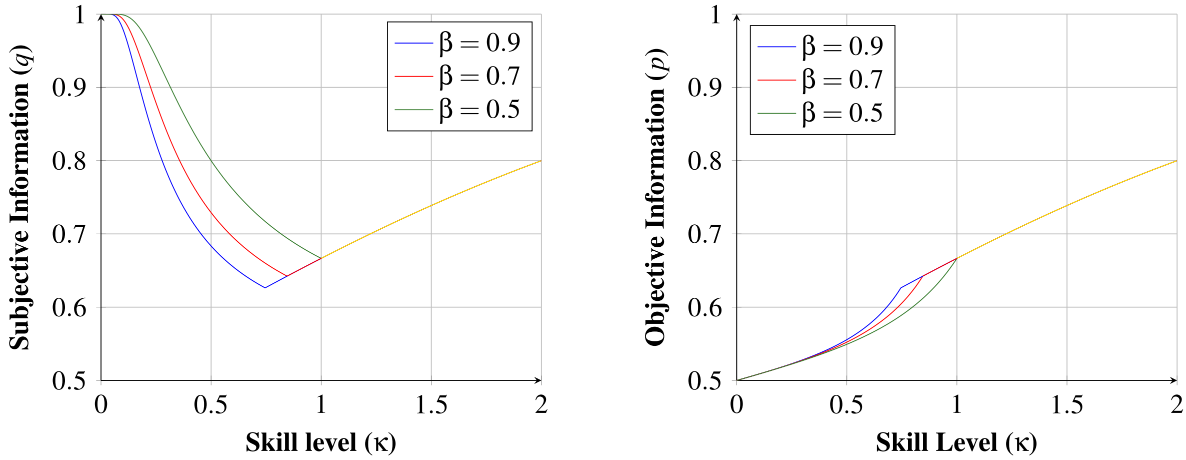
As specified above, assume that $\mu = 1/2$, $R = 1$ and $u(x) = x$, hence $EU_p = p$. After taking the first order condition, the result of this optimization problem is

$$\overline{p} = \frac{2^{\kappa}}{1+2^{\kappa}} = \frac{1}{1+2^{-\kappa}} \tag{2.19}$$

which is always larger than, or equal to $\mu = 1/2$, therefore the constraint never binds.

2.4 Implications of the Overconfidence Model

In this section I present and discuss the results of the Overconfidence Model with parameters $\mu = 1/2$, $\theta = 1/2$, $R = 1$, unit cost of information $\lambda(\kappa) = \frac{1}{\kappa}$, and unit cost of information



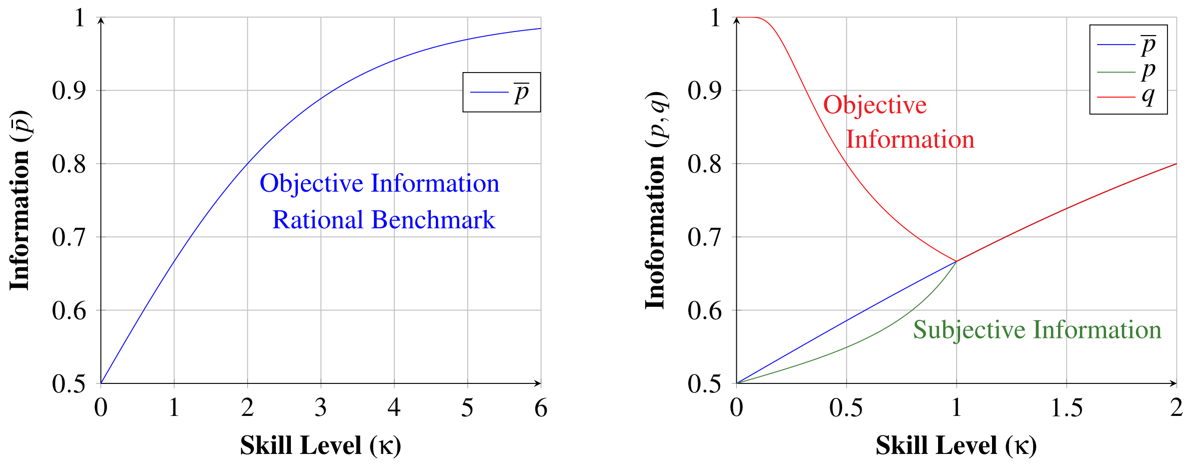
(a) Optimal Subjective Information (b) Optimal Objective Information

Figure 2.1: Optimal Objective and Subjective Information As a Function of Skill Level

distortion $\gamma(\kappa) = \beta\kappa$. The solution to the decision maker's problem under the Overconfidence Model under these assumptions is the objective information p defined by Equation (2.15) and subjective information q defined by Equation (2.16).

Figure 2.1 illustrates the optimal objective and subjective information for different awareness parameters as a function of the skill level. On the horizontal axis the increasing κ represents the learning process, and the different colors of the curves represent different awareness parameters. The left panel shows the evolution of the optimal subjective information. With $\kappa = 0$ skills, the decision maker chooses the highest possible level of subjective information $q = 1$, because at this point information distortion is free. As the skill level increases, the optimal amount of subjective information decreases until a point, where it turns back and starts to increase again. The right panel shows the evolution of the optimal objective information. With $\kappa = 0$ skills the decision maker cannot afford to gather any objective information and stays with her prior information, $p = \mu = 1/2$. As the skill level increases, the optimal objective information also increases, however there is a kink in its path.

The position and role of the kink of the objective information and the turning point of the subjective information we can understand more when we compare the Overconfidence Model



(a) No Information Distortion Model

(b) Overconfidence Model

Figure 2.2: Comparison of Optimal Objective and Subjective Information ($\beta = 0.5$)

to the benchmark No Information Distortion Model. Figure 2.2 illustrates the optimal objective and subjective information as the function of skill level for awareness parameter β=0.5. The left panel shows that under the No Information Distortion Model the optimal objective information monotonically increases as the skill level increases. There is no separate subjective information in this model, because it has to be the same as the objective information. In contrast, the right panel shows that the objective and subjective information can be different under the Overconfidence Model. The blue curve represents, as a benchmark, the result of the No Information Distortion Model. The green curve is the objective information under the Overconfidence Model. It increases as the skill level increases, but with a kink where the curve meets the optimal information under the No Information Distortion Model. The red curve is the subjective information, it initially decreases when the skill level increases. However, it turns back and increase after reaching the level of objective information. The three curves coincide after this point.

Overconfidence is the difference between the subjective and objective information. Under the Overconfidence Model, the gap exists for low levels of skills, and it decreases, then disappears as the skill level increases. This is the pattern described by the Dunning-Kruger effect.

The point where the overconfidence disappears is the point where the “No Underconfidence” constraint starts to be binding. Proposition 1. states these results formally.

Proposition 3. Under the Overconfidence Model:

1. The objective information equals the prior information when the skill level is zero. As the skill level increases, the objective information also increases.

2. The subjective information is the highest when the skill level is zero. As the skill level increases, the subjective information decreases until the “No Underconfidence” constraint is binding. After that point, the subjective information equals the objective information.

3. Overconfidence, defined as the gap between the subjective and the objective information, decreases as skill level increases.

Proposition 1. describes a path of subjective and objective information, which is in line with the observed pattern in the Dunning-Kruger effect literature [Kruger and Dunning, 2000, Sanchez and Dunning, 2018]. This is true, in general, for any positive awareness parameter, however, the exact shapes of these curves depend on the values of β. Figure 2.1 suggests that the increase of the objective information and the decrease of the subjective information is faster, and hence the decrease of overconfidence is faster, when β is higher. The same pattern can be observed in Figure 2.3, where the optimal objective and subjective information is illustrated for a wider range of β parameters, form 0.05 to 1. The smallest β parameters belong to the blue (subjective information) and green (objective information) curves on the right. The lower the awareness parameter is, the slower is the closing of the overconfidence gap. On the other hand, on the left the red and yellow curves represent the evolutions of the subjective and objective information when the awareness parameter is low. These curves quickly converge to the point

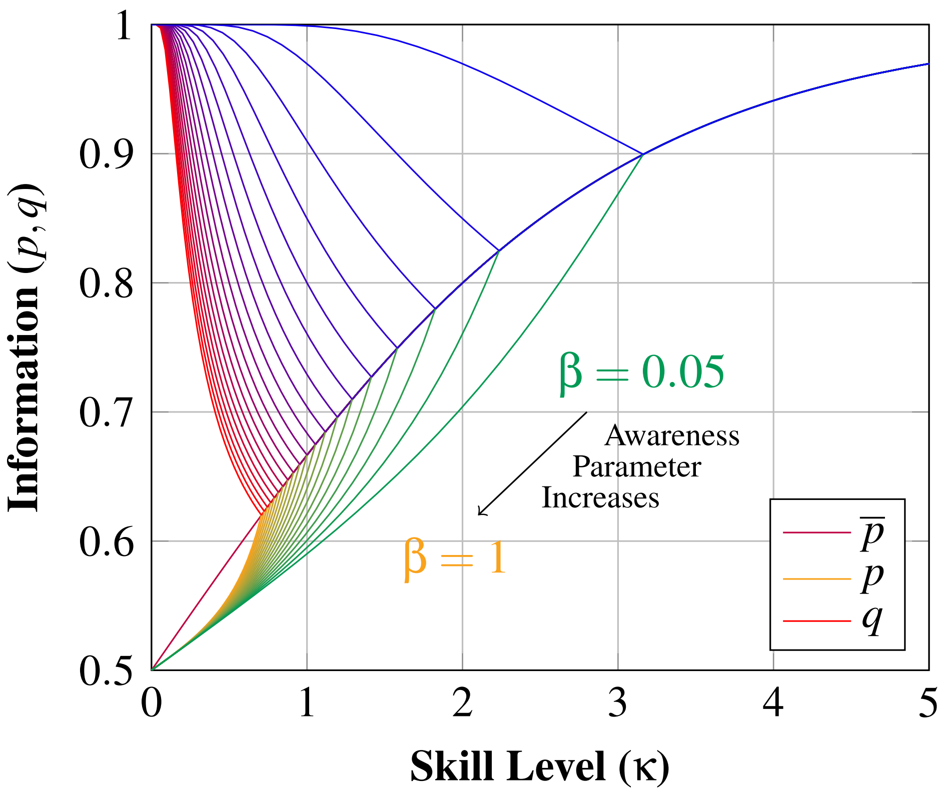


Figure 2.3: Optimal Objective and Subjective Information Under the Overconfidence and No Information Distortion Models (β from 0.05 to 1 from right to left)

where the overconfidence is dismissed. Therefore, the more aware the individual is about her bias, the faster she gets to correct it. Proposition 2 formalizes this argument.

Proposition 4. Under the Overconfidence Model, the higher the awareness parameter β is, the faster the objective information increases, and the faster the subjective information and overconfidence decreases.

Proposition 2. does not contradict the results of the psychology literature about Dunning-Kruger effect, however the above mentioned papers do not specifically examine the heterogeneity in individuals' awareness of their own overconfidence and the speed of their correction.

Finally, I compare the Overconfidence Model to the No Information Distortion Model based on the optimality of the information in a purely materialistic sense. Figure 2.3 reveals that overconfidence lowers the amount of objective information gathered, compared to the optimal amount of objective information in the No Information Distortion Model for a range of lower skill levels. It means that in terms of materialistic payoffs the decisions made under overconfidence are sub-optimal. Note that this result only holds as long as the subjective information is higher than the objective information, or in other words, when overconfidence is present. This is true when the skill level κ is lower than the threshold, where the No Underconfidence constraint starts to be binding. Proposition 3. summarizes these further results.

Proposition 5. In the skill domain where the decision maker is overconfident, the optimal objective information from the Overconfidence Model is lower than the optimal objective information from the No Information Distortion Model. That is, $p < \overline{p}$ whenever $p < q$.

Proposition 3. sheds light on the direct consequences of the overconfidence of the less skilled individuals. They choose to gather less than optimal objective information before solving a decision problem, meaning that they underestimate the effort they should optimally exert to solve the task. Hence they end up with lower than optimal expected outcomes in materialistic terms. Beyond this direct welfare-decreasing effect, an application in the next section highlights additional, more substantial issues induced by overconfidence. I show that it is an implication of the Overconfidence Model that the decision maker is not intrinsically motivated to learn when her initial skill level is too low.

2.5 Conclusions

In this project I presented a model that explains the inverse relationship between skills and overconfidence, a phenomenon called Dunning-Kruger effect in psychology. In my model, I assumed that both materialistic income and beliefs about future income provide utility. The decision maker is allowed to choose objective and subjective information separately in an entropy-based rational inattention model, where both the objective information and its distortion to subjective information have costs. The key assumption in the model is that while the cost of information decreases, the cost of information distortion increases with the increase of skill level. Therefore, the overconfidence of those with lower skills is the result of two opposite forces, the gain from overly optimistic beliefs and the cost of information distortion.

My model is also in line with former results in psychology. First, the model predicts the highest overconfidence for the least skilled individuals, and a decreasing trend in overconfidence as the skill level increases. This is the same pattern that was observed and called the Dunning-Kruger effect by Kruger and Dunning [2000].

Second, the implication that there is a threshold which separates the "low skill level" domain, where the decision maker does not see incentive to improve, and the "high skill level" domain, where the voluntary improvement starts is in line with the first two stages of the four stages of learning model by Noel Burch [Adams, 2017]. Namely, the low skill level domain can be interpreted as the "Unconsciously Unskilled" stage, and the high skill level domain as the "Consciously Unskilled" stage.

Finally, my model matches the key moments of the "beginners' bubble" hypothesis of Sanchez and Dunning [2018], the initial peak of overconfidence and its gradual decrease with experience. However, their description of the objective and subjective learning paths is more detailed. In their hypothesis, complete beginners start with low confidence, and then jumps quickly to their peak overconfidence after just a few trials. Then, after the gradually decreasing part which is supported by the current model as well, they predict that confidence levels, or probably start to increase again after a substantially high skill level has been reached. To build in these additional two feature to my model could be the subject of future research.

Chapter 3: Overconfidence, Excess Entry and Business Failure

3.1 Introduction

This chapter explains excess entry and high dropout rates of new firms utilizing the Overconfidence Model in Chapter 2. The setting specifies a threshold skill level above which entrepreneurs enter businesses, and a separate threshold skill level below which they drop out in the long run. I assume that the overconfident entrepreneur makes the entry decision based on the expected materialistic and belief utilities, net of information and information distortion costs. However, staying in the business in the long run only depends on the materialistic gains and its associated costs. The gap between the two thresholds characterize the skill levels of the entrepreneurs who do enter businesses, but unable to stay in. This application shows that overconfidence increases excess entry and high drop out rates in two ways. First, overconfident entrepreneurs enter businesses with the lack of adequate skill level to succeed. Second, even adequately skilled entrepreneurs may underperform and drop out due to their overconfidence.

Previous research suggests that overconfidence of entrepreneurs leads to excess entry to markets, which results in high probability of business failures. Camerer and Lovallo [1999] report that over 60 percent of entrants exit the market in less than five years, which they interpret as evidence of excess entry. They also provide experimental evidence that these excess entries are consequences of overconfident beliefs about one's own skills. I explain both the excess entries and the high percentage of business failures in this application of the Overconfidence Model from Chapter 2. Furthermore, I characterize the type of entrepreneurs who enter businesses but drop out in the long run, as well as those who prevail. I show that both entry decisions and drop outs are based on two individual attributes, the entrepreneurs' skill levels, and their preferences for overly optimistic beliefs.

In this setting I identify two separate conditions, one for the entry decisions of entrepreneurs, and one that specifies who is able to stay in business in the long run. For the entry decision, I take into account the net expected utility based on the Overconfidence Model, which consists of the materialistic utility, the belief utility, the cost of information and the cost of information distortion. However, in the condition for staying in business I consider only the materialistic gains and their information costs.

I show that there is a significant gap between the threshold skill level above which the entrepreneurs decide to enter businesses, and the threshold above which the firm can stay in business in the long run. This gap explains the high dropout rate reported in previous empirical studies, and shows that it is due to two different factors. First, there is excess entry because of the overly optimistic beliefs of the entrepreneurs about the probability of their future success. Second, the entrepreneurs' actual effort to succeed is below optimal because of their overconfidence. I compare these results to the behavior of the rational entrepreneurs in the No Information Distortion Model. I find that in their case there is no gap between the entry and dropout skill level thresholds.

When varying the awareness parameter, I find that stronger preferences for overly optimistic beliefs lead to lower skill level threshold in the entry decision and higher skill level threshold in the condition for staying in business. Therefore, the more overconfident entrepreneurs are more likely to enter with lower than the necessary skill level, and at the same time they are the ones who are more likely to drop out even when they would be skilled enough to perform adequately.

This application suggests that the market efficiently selects the higher performance in the long run, regardless of the overconfidence of the entrepreneurs. However, overconfidence can still result in some other form of inefficiency. Some of the entrepreneurs who drop out due to their higher preferences for overly optimistic beliefs have higher skill levels than some of those who can stay in business. These dropouts underperform, but if they could reduce their overconfidence, they could produce with less effort than those who stayed in business. Therefore, reducing overconfidence, if there is a way, is optimal from both the individuals' and from the society's point of view.

3.2 Setup

An entrepreneur chooses to enter a business where the outcome depends on effort and luck. Her outside option is a certain, fixed amount E. In case of entering the business, the entrepreneur's choices and outcomes are described based on the Overconfidence Model in Chapter 2. While in business, the entrepreneur faces a series of identical tasks, each of them results in either success, which pays off \$1, or in failure which pays nothing. The initial probability to succeed is μ, but the entrepreneur can choose to put in some effort to increase her chances. Moreover, she can also choose to form an overly optimistic belief about her probability of success.

Formally, the entrepreneur simultaneously chooses objective information p and subjective information q in a sequence of repeated tasks in business. The chosen objective and subjective information represent the probability of success in these attempts. p is the actual probability of success, and q is the entrepreneur's possibly overly optimistic belief about this probability. p and q are chosen in order to maximize a net utility function, which consists of a weighted average of a materialistic utility component $EU_p = p$ and a belief utility component $EU_q = q$, minus the cost of information $c(\mu, p)$ and the cost of information distortion $d(p,q)$.

In this application, I assume that the entrepreneur's choice of subjective information is defined as:

$$(p^*, q^*) = \arg\max_{p,q} \frac{1}{2}p + \frac{1}{2}q - c(\mu, p) - d(p,q) \tag{3.1}$$

$$s.t. \quad p \geq \mu \tag{3.2}$$

$$and \quad q \geq p \tag{3.3}$$

where

$$c(\mu, p) = \frac{1}{\kappa}(H(\mu) - H(p)) \tag{3.4}$$

$$d(p,q) = \beta\kappa(H(p) - H(q)). \tag{3.5}$$

The first two terms of the objective are the materialistic and belief utility components, with $\frac{1}{2} - \frac{1}{2}$ weights, respectively. The third term is the cost of information and the fourth term is the cost of information distortion. $H(p) = -p\log p - (1-p)\log(1-p)$ is the entropy of the binary distribution $(p, 1-p)$ over two states, success or failure. Parameter κ denotes the skill level of the entrepreneur, which in this setting will be a fixed, exogenous parameter that varies across individuals. Similarly, the awareness parameter β is also a fixed, exogenous individual characteristic. I utilize the variation of β in order to show how the entry decision depends on different tendencies to exhibit overconfidence, since lower β parameters results in more overconfidence.

I characterize the type of entrepreneurs who enter and who stay in business in two steps. First, I describe the condition for the entry decision as a function of the skill level. Second, I provide an additional constraint which determines the minimum skill level necessary to stay in business in the long run. Condition 1 states that an entrepreneur enters a business when the net expected profit – including materialistic and beliefs utilities, minus the costs of information and information distortion – is larger than the outside option, E.

Condition 1. Entry Decision

$$\frac{1}{2}p^* + \frac{1}{2}q^* - c(\mu, p^*) - d(p^*, q^*) \geq E$$

Condition 2 specifies that staying in business depends on whether the actual gains cover the actual costs, regardless of the additional utility from overly optimistic beliefs. The entrepreneur is forced to leave business in the long run, when her performance is constantly below the line where the business is more profitable than the outside option, that is when the objective information level p^* does not provide enough utility net of its information costs.

Condition 2. Stay in Business Constraint

$$p^* - c(\mu, p^*) \geq E$$

The explanation of different entry and exit criteria is the following. When fixed skill level is assumed, the net utility – including belief utilities and its costs – is also fixed, hence the agents would not exit business voluntarily despite materialistic losses. For example, the entrepreneur may interpret losses as temporary bad luck, while she can still see the future prospects positively. Reserves of the entrepreneur may cover these materialistic losses in the short run, however in the long run if the actual success rate p^* provides expected net materialistic utility lower than E, then the business will be forced to exit.

Note that this setting ensures that for the rational entrepreneur who acts based on the No Information Distortion Model the entry skill level threshold coincides with the long-term drop out skill level threshold, therefore these businesses are all sustainable in the long run. There could be business exits due to bad luck, for example, because of a series of negative draws, but these short term effects can be mitigated by having adequate emergency reserves to get through the negative periods. In contrast, if the entrepreneurs exhibit overconfidence, there is always a gap between the entry and the drop out skill level thresholds, which results in business failures in the long run, even when emergency reserves are available to mitigate losses due to temporary bad luck.

3.3 Implications

Figure 3.1 illustrates a numerical example with awareness parameter $\beta = 0.5$. The graph depicts net utilities as a function of skill level, κ. The blue line is the total net utility including materialistic utility, belief utility, the cost of information and the cost of information distortion. When the total net utility is above the outside option E, the entrepreneurs enter the market. This happens in two segments of the skill spectrum. The first segment consists of the very low skilled entrepreneurs below κ_0. Their actual success rate is low, and they leave the market early. The other segment is above the skill level κ_{entry}. Some entrepreneurs in this segment still perform below the threshold E, but their performance is not that low to drop out immediately. However, because of consistent underperformance, they are forced to leave business in the long run.

The red line shows the net materialistic utility minus information costs. A firm can stay in business in the long run when this utility is above the threshold E. This is the case when the skill level is above the threshold κ_{exit}. Those whose skill levels are between κ_{entry} and κ_{exit} do enter the market, but their actual performance is not enough to stay in business in the long run. The green line represents the net utility of the entrepreneur who does not exhibit any overconfidence, and acts based on the No Information Distortion Model. In this case the entry and exit skill level thresholds are the same. Compared to this benchmark case, the overconfident entrepreneurs have

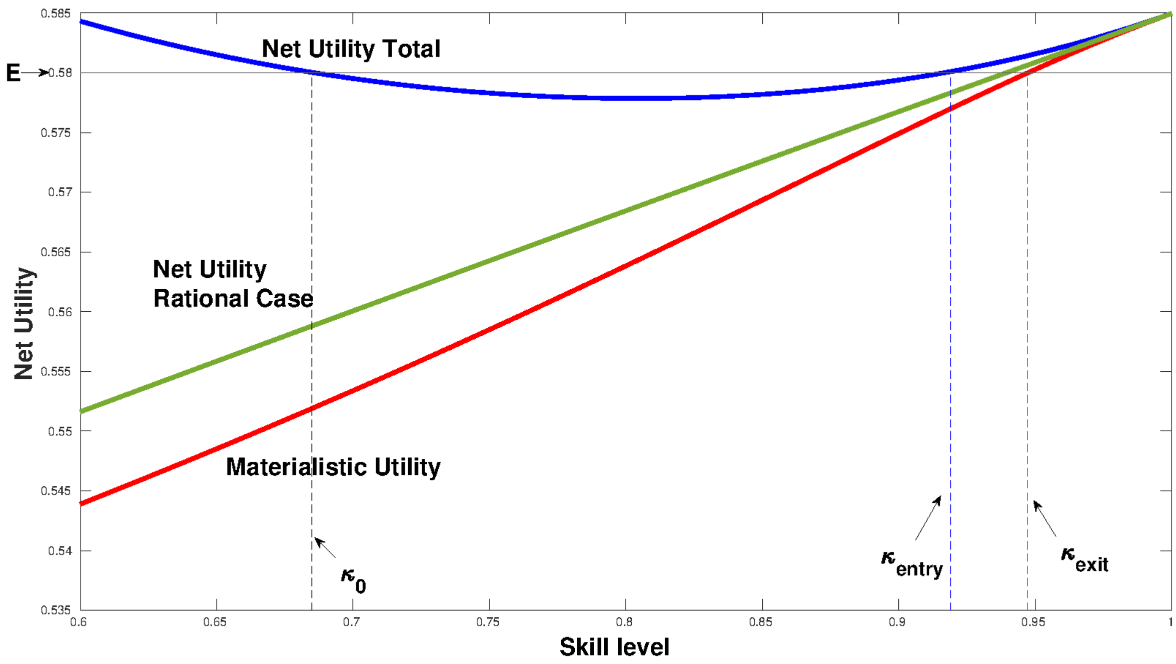


Figure 3.1: Entry and exit skill level thresholds $\beta = 0.5$

lower entry thresholds, but higher drop out thresholds. This result suggests that overconfidence causes both entries of entrepreneurs with too low skill levels, and also drop outs of entrepreneurs with high enough skill levels, but with high preferences for overly optimistic beliefs.

Figure 3.2 compares the entry and exit skill thresholds between two entrepreneurs who have different preferences for overconfident beliefs. These preferences are reflected in the model through the awareness parameter β. When the awareness parameter is higher ($\beta = 0.5$) then the individual is less overconfident, an when the awareness parameter is lower ($\beta = 0.48$), then the individual is more overconfident. The solid blue and red lines represent the net utility and the net materialistic utility for the less overconfident entrepreneur. Her entry and exit skill level thresholds are the same as before. The dash blue and red lines represent the net utilities for the more overconfident entrepreneur. When the overconfidence is higher, the entry skill level threshold moves lower and te exit skill level threshold moves higher. Therefore, the more overconfident

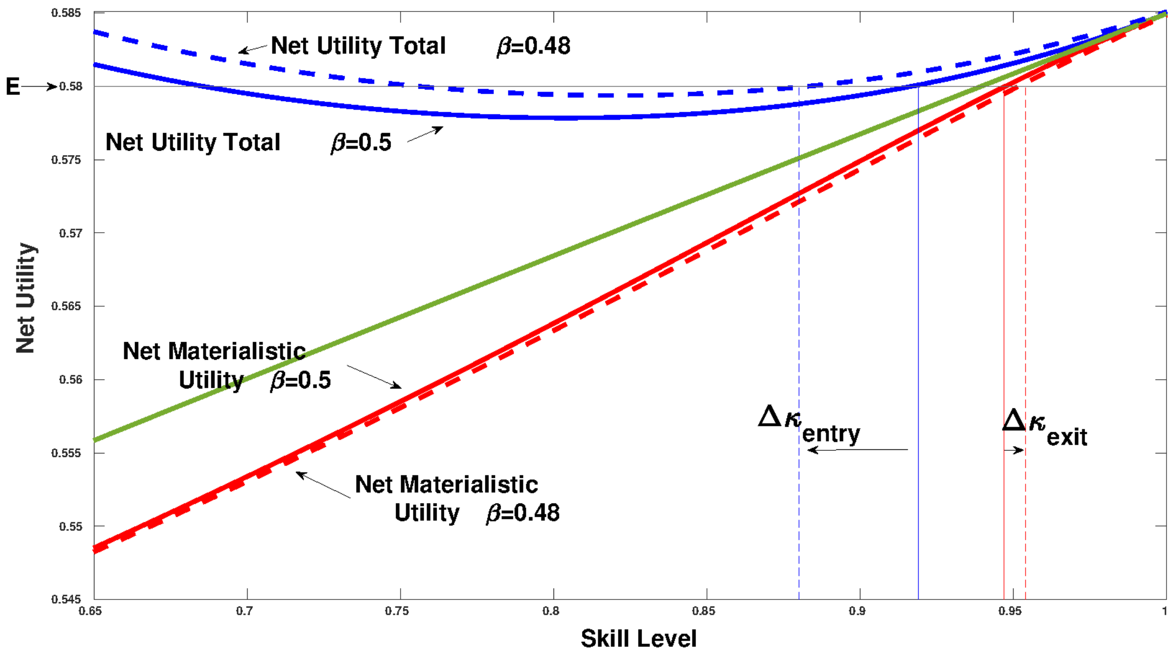


Figure 3.2: Entry and exit skill level thresholds $\beta = 0.5$ and $\beta = 0.48$

entrepreneurs are more likely to enter businesses with not high enough skill levels, and at the same time they are more likely to drop out even when they would have adequate skill levels to be successful, if they did not exhibit overconfidence.

Chapter A: Appendix to Chapter 1

A.1 Calculations and Proofs

A.1.1 Solution to the Buyer's Problem

The first order conditions can be taken for $f_B^p(A,v)$ and $f_B^p(R,v)$ separately. The first order condition with respect to $f_B^p(A,v)$ is:

$$U_B(p,v) - \lambda_B \left(log(f_B^p(A,v)) - log\left(q_S(v|p) \int_{v'} f_B^p(A,v)dv' \right) \right) - \theta(v) = 0 \tag{A.1}$$

where $\theta(v)$ is the Lagrangian multiplier. Solve this equation for $f_B^p(A,v)$:

$$f_B^p(A,v) = q_S(v|p)exp\left(\frac{1}{\lambda_B}U_B(p,v)\right)h(v) \tag{A.2}$$

where $h(v) = \int_{v'} f_B^p(A,v)dv' exp\left(-\frac{1}{\lambda_B}\theta(v)\right)$.

Similarly, first order condition with respect to $f_B^p(R,v)$:

$$-\lambda_B \left(log(f_B^p(R,v)) - log\left(q_S(v|p) \int_{v'} f_B^p(R,v)dv' \right) \right) - \theta(v) = 0 \tag{A.3}$$

Solve this equation for $f_B^p(R,v)$:

$$f_B^p(R,v) = q_S(v|p)h(v) \tag{A.4}$$

Plug in $f_B^p(A,v)$ and $f_B^p(R,v)$ to constraint (2): $q_S(v|p) = f_B^p(A,v) + f_B^p(A,v)$

$$q_S(v|p)exp\left(\frac{1}{\lambda_B}U_B(p,v)\right)h(v) + q_S(v|p)h(v) = q_S(v|p) \tag{A.5}$$

$$h(v) = \frac{1}{1 + exp\left(\frac{1}{\lambda_B}U_B(p,v)\right)} \tag{A.6}$$

Plug $h(v)$ back to Equations (6) and (7) gives the solution to the Buyer's problem. This result is summarized in Proposition 1.

A.1.2 Solution to the Seller's Problem

For the prices and values where $f_S(p,v) > 0$ and $g(v) > 0$, the first order condition with respect to $f_S(p,v)$ is:

$$U_S^E(p,v) - \lambda_s\left(log(f_S(p,v)) - log\left(g(v)\bar{f}_S(p)\right)\right) - \theta(v) = 0 \tag{A.7}$$

$$\Rightarrow f_S(p,v) = \bar{f}_S(p)exp\left(\frac{1}{\lambda_S}U_S^E(p,v)\right)g(v)h(v) \tag{A.8}$$

where $U_S^E(p,v) = \frac{U_S(p,v)}{1+exp\left(-\frac{1}{\lambda_B}U_B(p,v)\right)}$, $\bar{f}_S(p) = \int_{v'} f_S(p,v')dv'$, $h(v) = \exp\left(-\frac{1}{\lambda_B}\theta(v)\right)$, and $\theta(v)$ is the Lagrangian multiplier.

Denote $\mathcal{P}$ the support that contains all p prices of which the joint posterior distribution takes non-negative value. As it was discussed above, $\mathcal{P}$ is optimally chosen to be a discrete set of all possible prices in the continuous action space. Therefore, utilizing constraint (1.16) we can find $h(v)$:

$$g(v) = \sum_{p \in \mathcal{P}} \bar{f}_S(p) exp\left(\frac{1}{\lambda_S} U_S^E(p,v)\right) g(v)h(v) \tag{A.9}$$

$$\Rightarrow h(v) = \frac{g(v)}{\sum_p \bar{f}_S(p) exp\left(\frac{1}{\lambda_S} U_S^E(p,v)\right)} \tag{A.10}$$

Plugging this back in to the first order condition in Equation (A.8) gives the solution to the Seller's problem which is stated in Proposition 2.

A.2 Tables

λ_S	λ_B	#p	p_{min}	p_{max}	p_{av}
3	6	12	48.25	132.25	91.50
4	6	10	55.75	130.75	95.58
5	6	8	62.25	129.25	98.78
6	6	7	67.25	128.25	101.29
7	6	6	72.25	127.25	91.50
8	6	5	75.75	126.25	104.98
9	6	4	79.75	125.25	106.35

Table A.1: The number of price points (#p), the minimum price (p_{min}), the maximum price (p_max), and the average price (p_{av}) when $\lambda_B = 6$ fixed

λ_S	λ_B	#p	p_{min}	p_{max}	p_{av}
6	3	15	61.25	133.75	101.05
6	4	11	63.75	131.25	101.09
6	5	8	65.75	129.75	101.17
6	6	7	67.25	128.25	101.29
6	7	6	69.25	127.25	101.56
6	8	5	70.25	126.25	101.77

Table A.2: The number of price points (#p), the minimum price (p_{min}), the maximum price (p_max), and the average price (p_{av}) when $\lambda_S = 6$ fixed

λ_S	λ_B	I_S	C_S	I_B^{av}	C_B^{av}
3	6	0.8936	2.68	0.3432	2.06
4	6	0.6553	2.62	0.3847	2.31
5	6	0.4900	2.45	0.4100	2.46
6	6	0.3743	2.25	0.4247	2.55
7	6	0.2886	2.02	0.4335	2.6
8	6	0.2265	1.81	0.4383	2.63
9	6	0.1781	1.60	0.4408	2.64

Table A.3: The information flow of the Seller (I_S), the information cost of the Seller (C_S), the average information flow of the Buyer (I_B^{av}), and the average information cost of the Buyer (C_B^{av}) when $\lambda_B = 6$ fixed

λ_S	λ_B	I_S	C_S	I_B^{av}	C_B^{av}
6	3	0.5408	3.24	0.5365	1.61
6	4	0.4725	2.83	0.4954	1.98
6	5	0.4175	2.51	0.4584	2.29
6	6	0.3743	2.25	0.4247	2.55
6	7	0.3387	2.03	0.3942	2.76
6	8	0.3103	1.86	0.3666	2.93

Table A.4: The information flow of the Seller (I_S), the information cost of the Seller (C_S), the average information flow of the Buyer (I_B^{av}), and the average information cost of the Buyer (C_B^{av}) when $\lambda_S = 6$ fixed

λ_S	λ_B	$NetU$	$NetU_S$	$NetU_B$
3	6	12.10	9.88	2.21
4	6	11.15	9.12	2.03
5	6	10.46	8.55	1.91
6	6	9.95	8.12	1.83
7	6	9.56	7.79	1.77
8	6	9.27	7.53	1.74
9	6	9.05	7.33	1.71

Table A.5: Net Utility ($NetU$), Net Utility for the Seller ($NetU_S$) and Net Utility for the Buyer ($NetU_B$) under different unit information costs for the Buyer (λ_B) when $\lambda_B = 6$ is fixed

λ_S	λ_B	$NetU$	$NetU_S$	$NetU_B$
6	3	12.11	9.09	3.02
6	4	11.31	8.66	2.65
6	5	10.59	8.34	2.25
6	6	9.95	8.12	1.83
6	7	9.35	7.97	1.38
6	8	8.85	7.89	0.96

Table A.6: Net Utility ($NetU$), Net Utility for the Seller ($NetU_S$) and Net Utility for the Buyer ($NetU_B$) under different unit information costs for the Buyer (λ_B) when $\lambda_S = 6$ is fixed

λ_S	λ_B	$NetU$	$NetU_S$	$NetU_B$	I
7.5	4.4	10.41	7.96	2.45	0.8
7	4.9	10.24	7.99	2.25	0.8
6.5	5.4	10.10	8.05	2.05	0.8
6	6	9.95	8.12	1.83	0.8
5.5	6.7	9.75	8.19	1.56	0.8
5	7.4	9.61	8.31	1.30	0.8

Table A.7: Net Utility ($NetU$), Net Utility for the Seller ($NetU_S$), and Net Utility for the Buyer ($NetU_B$) when the total amount of information is fixed

A.3 Blahut-Arimoto Algorithm

I solve the Seller's problem using the Blahut-Arimoto algorithm (Stevens(2019), Blahut(1972)).

Parameters

λ_B: the Buyer's unit cost of information

λ_S: the Seller's unit cost of information

n: number of grids of the possible prices

n_v: number of grids of the value distribution

Initialize

$v(i), i = 1, \dots n_v$: array of possible values of v

$G(i) = \frac{1}{n_v}, i = 1, \dots n_v$: density function of value v

$p(j) = \frac{3}{2}\frac{j}{n}, j = 1, \dots n$: array of possible prices

$F(j) = \frac{1}{n}, i = 1, \dots n$: density function of prices

$d(i,j) = exp\left(\frac{1}{\lambda_S}\frac{p(j)-v(i)}{1+exp\left(-\frac{1}{\lambda_B}\left(\frac{3}{2}v(i)-p(j)\right)\right)}\right), i = 1, \dots n_v, j = 1, \dots n$: building blocks in the FOC

Iterate

$D = d * F^T$

$Z = (D.^{-1})^T .* G * d$

$F = F .* Z$

Further Results

$f(i,j) = F(j)\frac{d(i,j)}{D(j)}, i = 1, \dots n_v, j = 1, \dots n$: posterior beliefs of the Seller, joint distribution of v

and p_S

$f(:,p_s)$: posterior belief of the Seller given signal p_S

$f_A(:,p_S) = f(:,p_S)\frac{1}{1+exp\left(-\frac{1}{\lambda_B}\left(\frac{3}{2}v(i)-p_S\right)\right)}$: posterior belief of the Buyer given signal p_S and $Accept$

$f_R(:,p_S) = f(:,p_S)\frac{1}{1+exp\left(\frac{1}{\lambda_B}\left(\frac{3}{2}v(i)-p_S\right)\right)}$: posterior belief of the Buyer given signal p_S and $Reject$

$U_S = \sum_i \sum_j \frac{p(j)-v(i)}{1+exp\left(-\frac{1}{\lambda_B}\left(\frac{3}{2}v(i)-p(j)\right)\right)} f(i,j)$: expected utility of the Seller (before information costs)

$U_B = \sum_i \sum_j \frac{\frac{3}{2}v(i)-p(j)}{1+exp\left(-\frac{1}{\lambda_B}\left(\frac{3}{2}v(i)-p(j)\right)\right)} f(i,j)$: expected utility of the Buyer (before information costs)

Chapter B: Appendix to Chapter 2

B.1 Solution to the Overconfidence Model

When the constraint $q \geq p$ is not binding, the decision maker chooses q based on the first order condition:

$$(1-\theta)R - \beta\gamma(\kappa)\log q + \beta\gamma(\kappa)\log(1-q) = 0$$

The result of this optimization is

$$q^* = \frac{1}{1+2^{-\frac{(1-\theta)R}{\beta\gamma(\kappa)}}}.$$

Furthermore, the decision maker chooses p based on the first order condition:

$$\theta R - (\lambda(\kappa) - \beta\gamma(\kappa))\log p + (\lambda(\kappa) - \beta\gamma(\kappa))\log(1-p) = 0$$

The result of this optimization is

$$p^* = \frac{1}{1+2^{-\frac{\theta R}{\lambda(\kappa)-\beta\gamma(\kappa)}}}.$$

When the constraint is binding, then $p = q$ and the decision maker solves

$$\max_p EU_p - \gamma(\kappa)(H(\mu) - H(p))$$

$$s.t. \quad p \geq \mu$$

where $EU_p = R$. The result of this optimization is

$$\overline{p} = \frac{1}{1 + 2^{-\frac{R}{\lambda(\kappa)}}}$$

The constraint $q \geq p$ does not bind, if

$$\frac{1}{1 + 2^{-\frac{(1-\theta)R}{\beta\gamma(\kappa)}}} \geq \frac{1}{1 + 2^{-\frac{R}{\lambda(\kappa)}}}$$

$$\Rightarrow \frac{(1-\theta)\lambda(\kappa)}{\beta\gamma(\kappa)} \geq 1$$

Note that the second constraint, $p \geq \mu$ never binds if $\mu = 1/2$, because

$$\frac{1}{1 + 2^{-\frac{\theta R}{\lambda(\kappa) - \beta\gamma(\kappa)}}} \geq \frac{1}{2}$$

$$\Rightarrow \frac{\lambda(\kappa)}{\beta\gamma(\kappa)} \geq 1$$

which is always true, whenever $\frac{(1-\theta)\lambda(\kappa)}{\beta\gamma(\kappa)} \geq 1$.

In summary, the solution of the Overconfidence Model with "No Information Disposal" and "No Underconfidence" constraints is:

$$p = \begin{cases} \frac{1}{1+2^{-\frac{\theta R}{\lambda(\kappa)-\beta\gamma(\kappa)}}} = p^*, & \text{if } \frac{(1-\theta)\lambda(\kappa)}{\beta\gamma(\kappa)} \geq 1 \\ \frac{1}{1+2^{-\frac{R}{\lambda(\kappa)}}} = \overline{p}, & \text{if } \frac{(1-\theta)\lambda(\kappa)}{\beta\gamma(\kappa)} < 1 \end{cases}$$

$$q = \begin{cases} \frac{1}{1+2^{-\frac{(1-\theta)R}{\beta\gamma(\kappa)}}} = q^*, & \text{if } \frac{(1-\theta)\lambda(\kappa)}{\beta\gamma(\kappa)} \geq 1 \\ \frac{1}{1+2^{-\frac{R}{\lambda(\kappa)}}} = \overline{p}, & \text{if } \frac{(1-\theta)\lambda(\kappa)}{\beta\gamma(\kappa)} < 1 \end{cases}$$

References

George A. Akerlof. The market for ”lemons”: Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500, 1970. ISSN 00335533, 15314650. URL http://www.jstor.org/stable/1879431.

David M. Culter and Sarah J. Reber. Paying for health insurance: The trade-off between competition and adverse selection. *The Quarterly Journal of Economics*, 113(2):433–466, 1998. ISSN 00335533, 15314650. URL http://www.jstor.org/stable/2586909.

James H. Cardon and Igal Hendel. Asymmetric information in health insurance: Evidence from the national medical expenditure survey. *The RAND Journal of Economics*, 32(3):408–427, 2001. ISSN 07416261. URL http://www.jstor.org/stable/2696362.

David Powell and Dana Goldman. Disentangling moral hazard and adverse selection in private health insurance. *Journal of Econometrics*, 222(1, Part A):141–160, 2021. ISSN 0304-4076. doi: https://doi.org/10.1016/j.jeconom.2020.07.030. URL https://www.sciencedirect.com/science/article/pii/S0304407620302529. Annals Issue: Structural Econometrics Honoring Daniel McFadden.

B. Barsanetti and B. Camargo. Arbitrary bilateral bargaining in decentralized markets for lemons. Available at SSRN: https://ssrn.com/abstract=389322, 2021.

Igal Hendel and Alessandro Lizzeri. Adverse selection in durable goods markets. *American Economic Review*, 89(5):1097–1115, December 1999. doi: 10.1257/aer.89.5.1097. URL https://www.aeaweb.org/articles?id=10.1257/aer.89.5.1097.

Jean Tirole. Overcoming adverse selection: How public intervention can restore market functioning. *American Economic Review*, 102(1):29–59, February 2012. doi: 10.1257/aer.102.1.29. URL https://www.aeaweb.org/articles?id=10.1257/aer.102.1.29.

Thomas Philippon and Vasiliki Skreta. Optimal interventions in markets with adverse selection. *American Economic Review*, 102(1):1–28, February 2012. doi: 10.1257/aer.102.1.1. URL https://www.aeaweb.org/articles?id=10.1257/aer.102.1.1.

Christopher A. Sims. Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690, 2003a. ISSN 0304-3932. doi: https://doi.org/10.1016/S0304-3932(03)00029-1. URL https://www.sciencedirect.com/science/article/pii/S0304393203000291. Swiss National Bank/Study Center Gerzensee Conference on Monetary Policy under Incomplete Information.

Michael Woodford. Information-constrained state-dependent pricing. *Journal of Monetary Economics*, 56(S):100–124, 2009. doi: 10.1016/j.jmoneco.2009.06. URL https://ideas.repec.org/a/eee/moneco/v56y2009isps100-s124.html.

C. E. Shannon. Coding theorems for a discrete source with a fidelity criterion. *Institute of Radio Engineers National Convention Record*, 4:142–163, 1959.

G. E. Hoffer and M. D. Pratt. Used vehicles, lemons markets, and used car rules: Some empirical evidence. *Journal of Consumer Policy*, 10:409–414, 1987.

Eric Bond. A direct test of the ”lemons” model: The market for used pickup trucks. *American Economic Review*, 72(4):836–40, 1982. URL https://EconPapers.repec.org/RePEc:aea:aecrev:v:72:y:1982:i:4:p:836-40.

Eric W. Bond. Test of the lemons model: Reply. *The American Economic Review*, 74(4): 801–804, 1984. ISSN 00028282. URL `http://www.jstor.org/stable/1805152`.

Michael D. Pratt and George E. Hoffer. Test of the lemons model: Comment. *The American Economic Review*, 74(4):798–800, 1984. ISSN 00028282. URL `http://www.jstor.org/stable/1805151`.

M. D. Pratt and G. E. Hoffer. State used motor vehicle disclosure laws: Do they make a difference to the consumer? *Journal of Consumer Affairs*, 19:177–185, 1985.

M. R. Metzger. Cherries, lemons, and the ftc: Minimum quality standards in the retail used automobile industry. *Economic Inquiry*, 21:129–139, 1983.

MICHAEL D. Pratt and GEORGE E. Hoffer. The efficacy of state mandated minimum quality certification: The case of used vehicles. *Economic Inquiry*, 24(2):313–318, 1986. doi: https://doi.org/10.1111/j.1465-7295.1986.tb01812.x. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1465-7295.1986.tb01812.x`.

Simon Siegenthaler. Meet the lemons: An experiment on how cheap-talk overcomes adverse selection in decentralized markets. *Games and Economic Behavior*, 102:147–161, 2017. ISSN 0899-8256. doi: https://doi.org/10.1016/j.geb.2016.11.001. URL `https://www.sciencedirect.com/science/article/pii/S0899825616301300`.

Daniel Martin. Strategic pricing with rational inattention to quality. *Games and Economic Behavior*, 104:131–145, 2017. ISSN 0899-8256. doi: https://doi.org/10.1016/j.geb.2017.03.007. URL `https://www.sciencedirect.com/science/article/pii/S0899825617300519`.

Daniel Martin. Rational inattention in games: Experimental evidence. Available at SSRN: https://ssrn.com/abstract=2674224, 2016.

Doron Ravid. Ultimatum bargaining with rational inattention. *American Economic Review*, 110 (9):2948–63, September 2020. doi: 10.1257/aer.20170620. URL https://www.aeaweb.org/articles?id=10.1257/aer.20170620.

Junehyuk Jung, Jeong Ho (John) Kim, Filip Matějka, and Christopher A Sims. Discrete Actions in Information-Constrained Decision Problems. *The Review of Economic Studies*, 86(6):2643–2667, 03 2019. ISSN 0034-6527. doi: 10.1093/restud/rdz011. URL https://doi.org/10.1093/restud/rdz011.

Stephen L. Fix. Rate distortion functions for squared error distortion measures. Proceedings of the 16th Annual Allerton Conference on Communication and Control Computing (Oct., 1978) p. 704–711., 1978.

Filip Matějka. Rigid pricing and rationally inattentive consumer. *Journal of Economic Theory*, 158:656–678, 2015. ISSN 0022-0531. doi: https://doi.org/10.1016/j.jet.2015.01.021. URL https://www.sciencedirect.com/science/article/pii/S0022053115000320. Symposium on Information, Coordination, and Market Frictions.

Luminita Stevens. Coarse Pricing Policies. *The Review of Economic Studies*, 87(1):420–453, 07 2019. ISSN 0034-6527. doi: 10.1093/restud/rdz036. URL https://doi.org/10.1093/restud/rdz036.

R. Blahut. Computation of channel capacity and rate-distortion functions. *IEEE Transactions on Information Theory*, 18(4):460–473, 1972. doi: 10.1109/TIT.1972.1054855.

S. Arimoto. An algorithm for computing the capacity of arbitrary discrete memoryless channels. *IEEE Transactions on Information Theory*, 18(1):14–20, 1972. doi: 10.1109/TIT.1972.1054753.

K. Rose. A mapping approach to rate-distortion computation and analysis. *IEEE Trans. Inf. Theor.*, 40(6):1939–1952, September 1994. ISSN 0018-9448. doi: 10.1109/18.340468. URL https://doi.org/10.1109/18.340468.

Jonathan Levin. Information and the market for lemons. *RAND Journal of Economics*, 32(4): 657–66, 2001. URL https://EconPapers.repec.org/RePEc:rje:randje:v:32:y:2001:i:4:p:657-66.

Anke S. Kessler. Revisiting the lemons market. *International Economic Review*, 42(1):25–41, 2001. ISSN 00206598, 14682354. URL http://www.jstor.org/stable/2648717.

Anne-Katrin Roesler and Balázs Szentes. Buyer-optimal learning and monopoly pricing. *American Economic Review*, 107(7):2072–80, July 2017. doi: 10.1257/aer.20160145. URL https://www.aeaweb.org/articles?id=10.1257/aer.20160145.

D. Ravid, Anne-Katrin Roesler, and Balázs Szentes. Learning before trading: On the inefficiency of ignoring free information. *ERN: International Trade (Topic)*, 2019.

I. Csiszar. On the computation of rate-distortion functions (corresp.). *IEEE Transactions on Information Theory*, 20(1):122–124, 1974. doi: 10.1109/TIT.1974.1055146.

Ulrike Malmendier and Geoffrey Tate. Behavioral ceos: The role of managerial overconfidence. *Journal of Economic Perspectives*, 29(4):37–60, November 2015. doi: 10.1257/jep.29.4.37. URL https://www.aeaweb.org/articles?id=10.1257/jep.29.4.37.

Paul Heidhues, Botond Kőszegi, and Philipp Strack. Unrealistic expectations and misguided learning. *Econometrica*, 86(4):1159–1214, 2018. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/44955989.

David Dunning, Chip Heath, and Jerry M. Suls. Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5(3):69–106, 2004. doi: 10.1111/j.1529-1006.2004.00018.x. URL `https://doi.org/10.1111/j.1529-1006.2004.00018.x`. PMID: 26158995.

Colin Camerer and Dan Lovallo. Overconfidence and excess entry: An experimental approach. *American Economic Review*, 89(1):306–318, March 1999. doi: 10.1257/aer.89.1.306. URL `https://www.aeaweb.org/articles?id=10.1257/aer.89.1.306`.

Michael D. Grubb. Overconfident consumers in the marketplace. *Journal of Economic Perspectives*, 29(4):9–36, November 2015. doi: 10.1257/jep.29.4.9. URL `https://www.aeaweb.org/articles?id=10.1257/jep.29.4.9`.

Adam Harris and Ulrike Hahn. Unrealistic optimism about future life events: A cautionary note. *Psychological review*, 118:135–54, 11 2010. doi: 10.1037/a0020997.

Bertrand Russell. The triumph of stupidity. Mortals and Others: American Essays, vol. 2, p. 28, 1933.

Justin Kruger and David Dunning. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77:1121–34, 01 2000. doi: 10.1037//0022-3514.77.6.1121.

Carmen Sanchez and David Dunning. Overconfidence among beginners: Is a little learning a dangerous thing? *Journal of Personality and Social Psychology*, 114:10–28, 2018.

James Sawler. Economics 101-ism and the dunning-kruger effect: Reducing overconfidence among introductory macroeconomics students. *International Review of Economics Education*, 36:100208, 2021. ISSN 1477-3880. doi: https://doi.org/10.1016/j.iree.2020.100208. URL https://www.sciencedirect.com/science/article/pii/S1477388020300359.

David Dunning. Chapter five - the dunning–kruger effect: On being ignorant of one's own ignorance. volume 44 of *Advances in Experimental Social Psychology*, pages 247–296. Academic Press, 2011. doi: https://doi.org/10.1016/B978-0-12-385522-0.00005-6. URL https://www.sciencedirect.com/science/article/pii/B9780123855220000056.

Markus K. Brunnermeier and Jonathan A. Parker. Optimal expectations. *American Economic Review*, 95(4):1092–1118, September 2005. doi: 10.1257/0002828054825493. URL https://www.aeaweb.org/articles?id=10.1257/0002828054825493.

Roland Bénabou and Jean Tirole. Mindful economics: The production, consumption, and value of beliefs. *Journal of Economic Perspectives*, 30(3):141–64, September 2016. doi: 10.1257/jep.30.3.141. URL https://www.aeaweb.org/articles?id=10.1257/jep.30.3.141.

Bartosz Adam Mackowiak, Filip Matejka, and Mirko Wiederholt. Rational Inattention: A Review. CEPR Discussion Papers 15408, C.E.P.R. Discussion Papers, October 2020. URL https://ideas.repec.org/p/cpr/ceprdp/15408.html.

Markus M Mobius, Muriel Niederle, Paul Niehaus, and Tanya S Rosenblat. Managing self-confidence: Theory and experimental evidence. Working Paper 17014, National Bureau of Economic Research, May 2011. URL http://www.nber.org/papers/w17014.

David Eil and Justin M. Rao. The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3(2):114–38, May 2011. doi: 10.1257/mic.3.2.114. URL https://www.aeaweb.org/articles?id=10.1257/mic.3.2.114.

Don Moore and Paul Healy. The trouble with overconfidence. *Psychological review*, 115: 502–17, 05 2008. doi: 10.1037/0033-295X.115.2.502.

Don A. Moore and Derek Schatz. The three faces of overconfidence. *Social and Personality Psychology Compass*, 11(8):e12331, 2017. doi: https://doi.org/10.1111/spc3.12331. URL `https://compass.onlinelibrary.wiley.com/doi/abs/10.1111/spc3.12331`. e12331 SPCO-0805.R1.

Tali Sharot. The optimism bias. *Current Biology*, 21(23):R941–R945, 2011. ISSN 0960-9822. doi: https://doi.org/10.1016/j.cub.2011.10.030. URL `https://www.sciencedirect.com/science/article/pii/S0960982211011912`.

Brad M. Barber and Terrance Odean. Boys will be Boys: Gender, Overconfidence, and Common Stock Investment*. *The Quarterly Journal of Economics*, 116(1):261–292, 02 2001. ISSN 0033-5533. doi: 10.1162/003355301556400. URL `https://doi.org/10.1162/003355301556400`.

Gokul Bhandari and Richard Deaves. The demographics of overconfidence. *Journal of Behavioral Finance*, 7(1):5–11, 2006. doi: 10.1207/s15427579jpfm0701_2. URL `https://doi.org/10.1207/s15427579jpfm0701_2`.

Simon Gervais, J. HEATON, and Terrance Odean. Overconfidence, compensation contracts, and capital budgeting. *The Journal of Finance*, 66:1735 – 1777, 10 2011. doi: 10.1111/j.1540-6261.2011.01686.x.

Leonidas Enrique de la Rosa. Overconfidence and moral hazard. *Games and Economic Behavior*, 73(2):429–451, 2011. ISSN 0899-8256. doi: https://doi.org/10.1016/j.geb.2011.04.001. URL `https://www.sciencedirect.com/science/article/pii/S0899825611000698`.

Pietro Ortoleva and Erik Snowberg. Overconfidence in political behavior. *American Economic Review*, 105(2):504–35, February 2015. doi: 10.1257/aer.20130921. URL https://www.aeaweb.org/articles?id=10.1257/aer.20130921.

Roland Bénabou and Jean Tirole. Self-Confidence and Personal Motivation*. *The Quarterly Journal of Economics*, 117(3):871–915, 08 2002. ISSN 0033-5533. doi: 10.1162/003355302760193913. URL https://doi.org/10.1162/003355302760193913.

Andrew Caplin and John Leahy. Psychological expected utility theory and anticipatory feelings. *The Quarterly Journal of Economics*, 116(1):55–79, 2001. ISSN 00335533, 15314650. URL http://www.jstor.org/stable/2696443.

Joachim Krueger and Ross Oakes Mueller. Unskilled, unaware, or both? the better-than-average heuristic and statistical regression predict errors in estimates of own performance. *Journal of Personality and Social Psychology*, 82:180–188, 01 2002. doi: 10.1037//0022-3514.82.2.180-188.

Justin Kruger and David Dunning. Unskilled and unaware - but why? a reply to krueger and mueller (2002). *Journal of personality and social psychology*, 82:189–92, 03 2002. doi: 10.1037//0022-3514.82.2.189.

Linda Adams. ILearning a skill is easier said than done. https://www.gordontraining.com/free-workplace-articles/learning-a-new-skill-is-easier-said-than-done/, 2017.

Xavier Gabaix. Chapter 4 - behavioral inattention. In B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, editors, *Handbook of Behavioral Economics - Foundations and Applications 2*, volume 2 of *Handbook of Behavioral Economics: Applications and Foundations 1*, pages 261–343. North-Holland, 2019. doi: https://doi.org/10.1016/bs.hesbe.2018.11.001. URL https://www.sciencedirect.com/science/article/pii/S2352239918300216.

Russell Golman, David Hagmann, and George Loewenstein. Information avoidance. *Journal of Economic Literature*, 55(1):96–135, March 2017. doi: 10.1257/jel.20151245. URL `https://www.aeaweb.org/articles?id=10.1257/jel.20151245`.

Christopher A. Sims. Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690, 2003b. ISSN 0304-3932. doi: https://doi.org/10.1016/S0304-3932(03)00029-1. URL `https://www.sciencedirect.com/science/article/pii/S0304393203000291`. Swiss National Bank/Study Center Gerzensee Conference on Monetary Policy under Incomplete Information.

Filip Matejka and Alisdair McKay. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105, 06 2011. doi: 10.2139/ssrn.1886964.

Andrew Caplin and Mark Dean. Revealed preference, rational inattention, and costly information acquisition. *American Economic Review*, 105(7):2183–2203, July 2015. doi: 10.1257/aer.20140117. URL `https://www.aeaweb.org/articles?id=10.1257/aer.20140117`.

Andrew Caplin, Mark Dean, and John Leahy. Rationally inattentive behavior: Characterizing and generalizing shannon entropy. Working Paper 23652, National Bureau of Economic Research, August 2017. URL `http://www.nber.org/papers/w23652`.

Mark Dean and Nate Leigh Neligh. Experimental tests of rational inattention. 2017.

Michaela Pagel. A news-utility theory for inattention and delegation in portfolio choice. *Econometrica*, 86(2):491–522, 2018. ISSN 00129682, 14680262. URL `http://www.jstor.org/stable/44955975`.

Theodore Paul Wright. Factors affecting the cost of airplanes. *Journal of the Aeronautical Sciences*, 3:122–128, 1936.