An Attempt at a literature review of "The Market for Lemons" Theory and a Case Study in Algerian Local E-commerce

Salmi, Mabrouka

National School of Statistics and Applied Economics

15 March 2022

An Attempt at a literature review of "The Market for Lemons" Theory and a Case Study in Algerian Local E-commerce

Mabrouka Salmi[1]

[1] National School of Statistics and Applied Economics (ENSSEA), Koléa-Tipaza, Algeria
salmi.mabrouka@enssea.net

Abstract. This paper attempts to understand the theory of "the market for lemons" proposed by the American economist Akerlof in 1970, review some related works, and apply this theory in an Algerian context. The theory of Akerlof is key to the information economy. One of the first pieces of research that defined adverse selection is when the two parties in a transaction have different levels of information. We present our example, not well established in the literature, of asymmetrical information in the local e-commerce market, which explains the e-sellers problem with e-customers behaviors. We describe the impact of such issues, such as the loss of e-sellers caused by behaviors of e-customers, and propose solutions based on the literature. Finally, we are beyond reviewing such landmark research or adding anything to it, related works are not fully discovered, and we present a humble piece that may exhibit any flaws.

Keywords: Information Asymmetry, Adverse Selection, Lemons Market, Local E-commerce.

1 Introduction

In any dealing, if one side is well informed yet perfectly informed about the merchandise quality (transaction) than the other side, we are in a case of asymmetrical information. Asymmetry of information leads to adverse selection. Adverse selection is when one party to a transaction has more or better information than the other party, which can cause the parties to make different choices and inefficient outcomes. An example of adverse selection would be when you are looking for car insurance, and you know more about your driving habits than your potential insurer does. Another example of asymmetrical information is when the insured, in case of a car accident, presents misleading information to the insurer to benefit some financial gains; it is known as car insurance fraud. In a market, potential buyers have asymmetrical information about the quality of a product. On the one hand, sellers know how good their product is, and on the other hand, buyers know only how much it costs and its features.

Adverse selection is the idea that you will be less likely to buy it when you know more about the product than the seller. As Akerlof (Akerlof, 1970) has provided an example in health insurance, people more than 65 years old can more easily assess the risks involved than the insurance company.

Adverse selection and asymmetric information have been a hot topic in research. Akerlof's research illustrated the lemons model based on the example of the used cars market and extended his demonstration with examples in the money market and labor market. Besides the effects of quality uncertainty in terms of reductions in goods quality and size of the market. Akerlof has taken the used automobile market, and it was the case for numerous researchers who further developed the lemons model and added to it (Wilson, 1980) and (Kessler, 2001). While Wilson proved the existence of multiple equilibria in the lemons model, Kessler proposed

various forms of information structures in adverse selection in markets. Their research findings are stated in the related work.

Our objective was to attempt to review Akerlof's research, but instead, the essentials of his work, like the lemons model. Further, we try to overview some research that developed Akerlof's conclusions. Finally, we attempt to apply what we have understood in an example of information asymmetry existing in local e-commerce in Algeria.

The rest of this paper is organized as follows: Section 2 explains the proposed methodology. Related works and the article of Akerlof are described in section 3, the example of application is presented in section 4, and section 5 concludes this work.

2 Methodology

In this section, we describe the steps that we have followed to achieve the objectives of this work.

2.1 Search strategy

The lemons theory is well-known, and it has been attracting researchers from various fields up to this point in time. It was proposed by the American economist Akerlof, who gained a Nobel Price shared with two other economists regarding his contributions to the economy of information. We observe that research on this topic is extensive, to the extent that it cannot be thoroughly reviewed in this report. However, this has been noticed after preliminary research conducted in the Google Scholar bibliographic database using relevant search terms (keywords) combined with logical operators such as 'AND' and 'OR' to extract all the published research on the topic. Due to the over-explore of this topic, we tend only to search the keywords in the titles of research articles. The data extraction results with and without this option are stated in table (1). Figure (1) shows the selected keywords and how they were used.

Based on this structured search strategy, we select only relevant research on "the market for lemons" theory after following other steps like scanning and filtering several collected papers through the abstract and the number of citations. We understood the topic that allowed us to select some research on the lemons theory, which is more than five articles in a time range of [1970,2021]. First, we review the initial work of Akerlof and further some works that aimed to explain and add to the lemons market theory. Second, we present an applications taxonomy where we review the theory applications in various fields. The selection of papers to conduct this part is non-random. We pick some papers that verify the inclusion criteria and then select papers working in different fields. To explore as many applications field as possible, we focus on discovering the different fields where the lemons theory has been applied rather than reviewing the work in each field.

Fig. 1. Search keywords used for data extraction

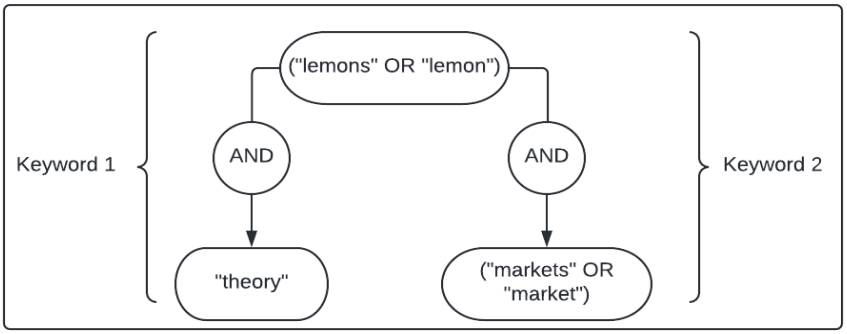


Table 1. Search keywords results using Google Scholar. Keywords 1 and 2 are mentioned in figure 1

Search terms	Advanced search options and the results of each		Date and time of search execution
	Search in all the document	Search in title	
Keyword 1	266000	41	00:12, 22/02/2022
Keyword 2	240000	651	00:11, 22/02/2022

2.2 Application example

To apply this theory, we took an example from the Algerian economy, local e-commerce. We noticed that this market involves some characteristics of information asymmetry. However, the issue we present is rarely covered in research due to its specificity to the local Algerian e-commerce. We present the statistics of a local e-commerce company in Algeria and open a discussion on how to adapt the model of Akerlof to it if possible.

3 Literature Review

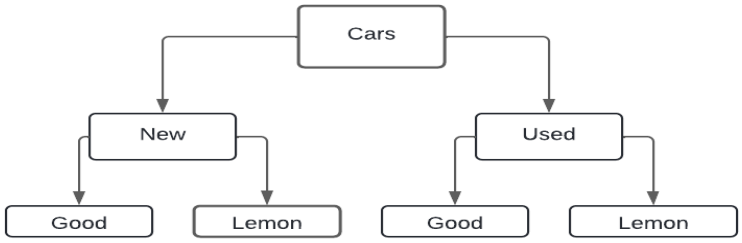
3.1 Understanding The Lemons Market theory

To better understand the theory of the lemons market we reproduce the driven objectives as well the lemons model and the example of the original paper (Akerlof, 1970). Akerlof aimed to explain that differences in the quality of goods existing in the market yield fundamental problems for the market concepts. Besides, the association of differences in quality uncertainty may clarify important institutions in the job market.

In his paper, he also provided a structure to the difficulty of the underdeveloped countries' businesses, and mainly he structured how to determine the economic costs of dishonesty. In addition, he applied his theory to the structures of money markets and the labor market.

Akerlof's model with Automobile (Akerlof, 1970)**.** As an example, figure (2) depicts the assumptions that Akerlof has used as a basis to illustrate his thoughts. It is given: buyers of a new automobile in this market don't know the quality of the car which can be good or "lemon", but these assumptions are made: $(1-q)$: the probability that it is a lemon, so q: the proportion of the produced good cars, and $(1-q)$: the proportion of lemons.

Fig. 2. Assumptions of Akerlof concerning the cars



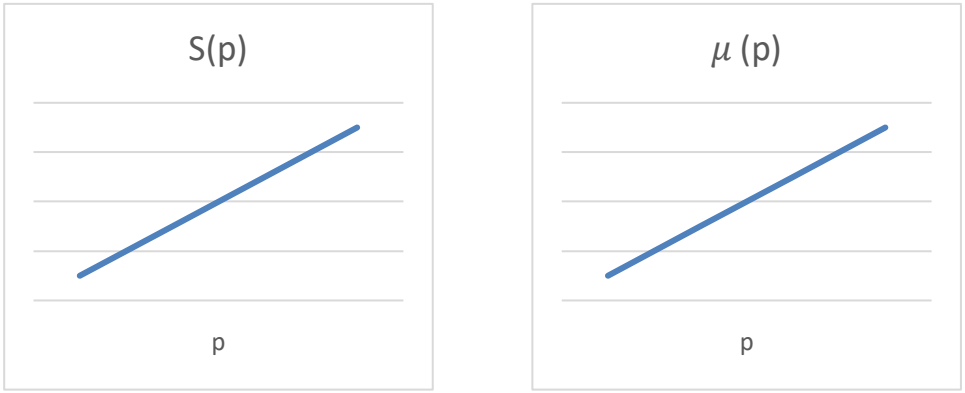
The buyer's experience of his car allows assigning a new probability to the event that his car is a lemon which is more accurate than the initial estimate (original). What happened is called by Akerlof "Asymmetry is available information": only sellers currently know the quality of a car more than the buyers. At the same time, both good and bad cars (lemons) must be sold at the same price. Buyers cannot distinguish between good cars and bad cars. The valuation of types of cars is different. However, if they have the same valuation, trading a lemon at the price of a new car is advantageous. The probability q that a bought car is good augments, and its probability of being bad decreases. As a result, an owner of a good automobile can receive neither the true value of his car nor the expected value of a new car.

Gresham's law. (Akerlof, 1970) In the market (in Akerlof's example), most traded cars will be bad (lemons), and good cars may be severely less traded or not at all. Akerlof noted that Gresham's law, which has been modified, presents an instructive analogy, but it is not complete enough. Because per Gresham's law, the reason behind driving out the good cars by bad cars is that they sell at the same price, which makes no difference for buyers between good and lemons, and only the sellers can differentiate them. Here it is presumable the availability of the information about the quality of the merchandise.

Asymmetrical information. (Akerlof, 1970) Akerlof extended his example of bad and good cars to a market containing different grades of goods in terms of quality, on a scale: bad, not so bad, medium, not so good, and good. The driving out process happens between goods of two successive quality grades, starting with bad goods such as those driving out the not so bad ones. In addition, he insisted that no market can exist in the presence of such a sequence of driving out events.

They assumed that the demand for used automobiles primarily depends on two variables. p: the price of the automobile; μ: the average quality of used cars traded. Therefore, $Q^d = D(p, \mu)$. Moreover, S: the supply of used cars and the average quality μ are positively correlated with the price p as shown in figure (3). Per Akerlof's statement, as the price decreases, the quality will decrease, and possibly no goods will be marketed at any price level. Where in equilibrium, the supply and the demand are equal for a certain average quality: $S(p) = D(p, \mu(p))$.

Fig. 3. The curve of the supply S and the quality of the goods μ in function of the price p, the plots show the positive linear correlation between each of them and the price.



Utility function. (Lemon model) Akerlof gave a further example extracted from utility theory. Assuming the existence of only two groups of traders (one and two). Where $\mathcal{U}_i$ is the utility of the group i, the utility function for each group is well stated in the original paper (pg.490) (to avoid repetition).

Reasons behind the linear utility function. Akerlof explained his choice of a linear utility function, first to avoid algebraic complication, second to concentrate on the effects of asymmetry of information based on concave utility function besides the risk variance effects of uncertainty and special effects which he discussed, third to keep the same incrementation of the utility whether by the first car, the second car, …, the k^{th} car, which refers to the odd characteristics of $\mathcal{U}_1$and $\mathcal{U}_2$.

Lemons model. in the example proposed in (Akerlof, 1970), it is also assumed that traders have an identical type which is "von Neuman-Morgenstern" maximizers of expected utility, that group one has N cars with quality x of $U_{[0,2]}$ distribution, group two has no cars. The price is set to unity of the goods in M. Y_i: the income of all type i traders.

The demand for used cars is equal to the sum of D_i, which is the demand of the i^{th} group. By ignoring indivisibilities, type i traders' demand for automobiles is described in the paper (pg.491) as well as the total demand $D(\mu, p)$

Symmetric Information. Assigning x: the quality of all cars follows $U_{[0,2]}$, the formulas of the demand and supply in symmetric information, can be found in the original paper (pg.492) and the three cases of equilibrium that can result in.

Akerlof mentioned that slight changes are included in the three cases of equilibrium if the same probabilistic estimates provide the quality of individual automobiles of both traders with p: the expected price.

Dishonesty's Cost. According to Akerlof, in a market with the lemons model, the cost of dishonesty corresponds to dishonest trades, which tend to drive out honest (legitimate) trades. It consists of the amount by which buyers are cheated, which the sellers set, and the loss created from driving honest businesses out of existence.

3.2 Related Works

Wilson proposed the existence of multiple equilibria (Wilson, 1980) in the market with significant adverse selection that can be sorted in an ascending order using the Pareto criterion. He investigated particularly the dependence of the equilibria on how the price is decided ("mechanism"). He experienced three mechanisms to set the price, whether by the sellers, buyers, or the auctioneer. A single equilibrium was only in the third mechanism. When the buyers or sellers decide the price, an excess in supply may be caused by the price range; we refer to his paper for his experiments.

Bond tested the lemons on used pickup trucks (Bond, 1982), where the maintenance quantity necessary on a truck is used as the quality metric. A lemon truck needs much more maintenance than a standard truck. His study revealed no considerable difference between new trucks and used trucks in maintenance. As a result, the hypothesis that bad products evicted the good ones is rejected, which means it is not a lemons market. Bond explained his findings with the development of counteracting institutions like seller's reputation and warranties which tend to balance the information of buyers and sellers. These results opened the debate on whether or not lemon markets exist in the used vehicles market, highlighted and reviewed by Hoffer (Hoffer and Pratt, 1987).

Kessler used the example of used cars of Akerlof and studied various structures of information in a market of lemons (Kessler, 2001). He divided the sellers into two groups, the first is perfectly informed about the machine quality, and the second is imperfectly informed (average to less). Moreover, he compared the performance of the market under the two aforementioned informational structures. It is concluded that uniformed sellers compete with a relatively higher price than the case of informed sellers, which boosts the average quality of the market, a less informed seller's reservation price is greater than that of a seller who perfectly knows his car.

Quality uncertainty in used car markets remains a critical issue. (Zavolokina, Schlegel and Schwabe, 2021) A group of 'car-related ecosystem' organizations launched an approach to coping with it, the " Carddossier " project, in 2019. It is a Swiss blockchain-based electronic platform that works to preserve and share all information about a car's history from manufacturers to consumers. The concept is to reduce asymmetrical information by recording all existing information about the car's lifecycle, thereby removing uncertainty in quality and maintaining buyer-seller trust. Several research studies improved different aspects of this blockchain platform, such as governance and data privacy. This paper discussed how to reduce the information asymmetries by improving the platform's design.

Some examples of counteracting proposed by Akerlof aim to reduce the impact of quality uncertainty, such as brand names that reflect the quality and the existence of 'retaliation' if the quality is not as expected by the consumer. Brand names also exist in the education and labor market, where "the practice of licensing" by which the majority of skilled labor have specified certification (Ph.D., Bachelor, etc.) implies their level of proficiency. (Ryoo, 1996) focused on asymmetric information in a specific type of problem in the labor market, the professional service market involving services like law, medicine, and others. Where the service quality is hard to be identified even after purchase, he claimed that the quality of service is similar to a worker's quality or talent. He developed the lemons models considering the features of the professional services markets, in terms of supply elasticity in the long term, wide ranges of prices, and large service providers (such as hospitals). His research concludes that professional certifications are

unlikely to be a long-term remedy to the adverse selection issue, as claimed before by Leland (1979) (Leland, 1979).

Taxonomy of Application Markets. A variety of markets have been subject to information asymmetry research, based on the lemons model and adapted models developed to handle adverse selection in markets considering the specificities of each one. In table (2), we present the idea of information asymmetry in each research from each field. We refer to the research papers for further information. The papers in the table have been selected through the abstract and title scanning only.

Table 2. Taxonomy of application markets in information asymmetry research

Market	Research works	Short summary
Pharmaceuticals	(Light and Lexchin, 2021)	In the pharmaceuticals market, some 'clinicians' have a monopoly on deciding prescribed drugs and information told to patients, which frame medical diagnosis and prescribing decisions. This research examination defined how firms tend to reduce knowledge about the hidden side effects and harms in new drugs… which shows pharmaceuticals as a lemon market.
Tourism	(Baggio and Baggio, 2011)	In the tourism market, as the authors explained, travelers, even with the arrival of the internet, still cannot comprehensively analyze the quality of their considered distention before booking a trip, which leads to quality uncertainty about offered travel services, destinations, and companies…
E-commerce	(Lin, 2012)	According to Lin's research, the internet provides buyers with robust search tools to obtain a large amount of information. Consequently, the buyer tends to be overwhelmed by the abundance of options, complicating their decision-making. However, it is noted that the problem is the unavailability of quality information which increases the quality uncertainty…

4 Application Example in the Algerian Example

Adverse selection is a type of market structure where the buyers and sellers have different information, leading to the buyers being mistreated. In this type of market, there are two types of participants: informed and uninformed. The informed group knows more about what they are buying than the uninformed group, which creates an imbalance in the marketplace because the informed group can take advantage of the uninformed group by charging higher prices for lower-quality goods. For example, if a person needs to buy a car but does not know much

about cars and has to rely on the seller's information, they are at an advantage because they can manipulate or withhold information from the buyer.

4.1 Local e-commerce in Algeria problem with customer's behavior: a case study

With the prevalence of the internet, another market appeared: the e-market. The E-market is quite different from the real market; sellers and buyers use the internet to interact. E-sellers have a marketplace where their products are exposed, and E-buyers use this marketplace to purchase. Further, each product is usually perfectly described, and information is available about the products for buyers, and they can ask for further information and find answers to their inquiries.

Purchase in these online markets is performed through a set of procedures. This example focuses on local e-commerce in Algeria, which involves specific procedures for Algeria resident people, which can change depending on the local e-commerce company itself. In the following, we present a common problem faced by such companies with Algerian e-customers behaviors in the purchase. Moreover, we take the case of anonymous local e-commerce and present the problem using real-world data (its data). Local e-commerce companies in Algeria use different procedures of purchase. For that, this case we present can be slightly changed depending on regularities set by the company.

Based on our understanding of information asymmetry, we propose that it is a case of an asymmetrical information issue. Moreover, the local e-commerce market involves asymmetrical information between e-sellers and e-buyers.

Case study. Let's name the local e-commerce company $'X'$; X has a marketplace where buyers can consult the products and select products to purchase. Where the products are fully described (with texts, real-world pictures, and videos), any needed information about any product on the marketplace is available. We assume that the e-buyers are well informed about the product, its quality, and its price.

Purchase procedures. We present the steps of purchase according to this company as well as the costs involved in this process:

- The first step is *ordering*, by which the buyer makes an order through the platform, this cannot be done without Marketing efforts carried out by the company X; each buyer who enters the platform and orders a product already costs the company marketing costs, which in our case is equivalent to 300 DA, here we set m is the e-marketing cost per customer who orders a product. Here, the e-buyer must enter necessary information like the name, address, phone number, and the product to buy to complete his order. This data is well stored in the platform and services to the next steps.
- The second step is *order confirmation*, The company, at this stage, confirms the order of each buyer using the platform database. Through a phone call with buyers, the company assures their knowledge about the product, its price, the purchase process, and if they can carry on their orders. Here many issues pop up first; some buyers who passed the first step are unavailable for phone calls (wrong phone number, no response, or completely unavailable). The company repeats calling many times to confirm their orders.
  Consequently, three events happen:
  - The buyer responds to one of the phone calls and confirms.

- The buyer responds to one of the phone calls and does not confirm (long or short phone call).
- The buyer is not reached by phone after many attempts.

So, whether orders are confirmed or not, costs of phone calls are already made. In our example, the phone calls costs depend on time (t). We will assume a constant cost of phone calls, regardless of how many times calls are made, and c is the phone calls cost per buyer. After confirmation, the local company assures the correct address and the delivery time.

— The third step is the delivery of the product, X: the e-seller prepares orders with packaging and the shipping information appearing on it. In addition, it takes them to the delivery company to occupy with delivering the product to the buyer in a well-known time by both sides. Here some costs are involve packaging and transporting the products to the delivery. We call them d. The delivery company calls the buyer at the specified time, determined in the second step. The delivery process is that the delivery man only calls the buyer three times (attempts). The buyer responds to one of them, takes his product from the address he has already chosen, and pays the purchased product with price p. Alternatively, the buyer responds to the delivery man and changes the time of the delivery time to later (all the processes will be repeated on that later time). Alternatively, the buyer does not respond within the three times the product, which returns the product to the e-seller. Alternatively, the buyer responds and cancels his order, and the product returns to the e-seller. In the last two cases, the cost of returning the products to the e-seller by the delivery company is involved. It is equal to 350 DA. We note it r. we assume the e-seller, in case of product return, will only take charge of this cost r.

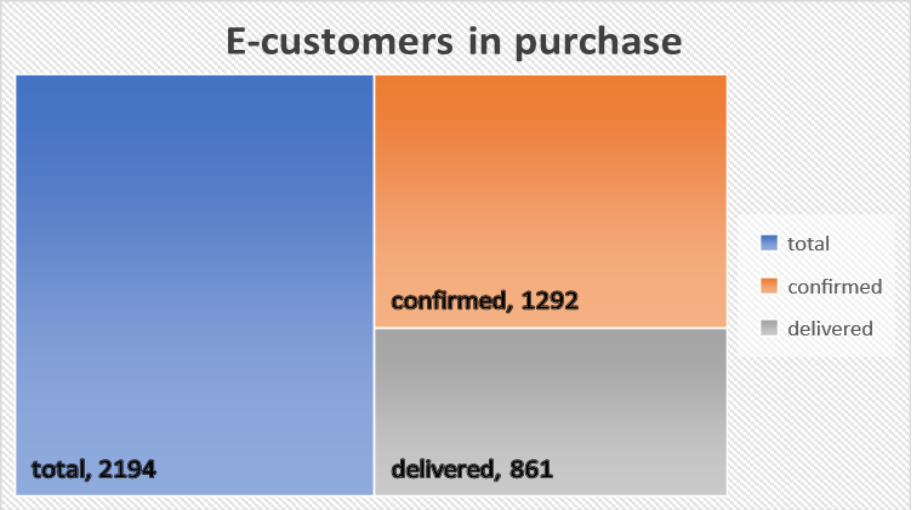
Statistics of the local e-commerce company. We describe the data of X through a treemap. The data consists of

- 'orders': the number of orders made through the platform means the e-buyers satisfied step one in the purchase process
- 'confirmed': the number of orders confirmed refers to buyers confirming their interest in the purchase and passing the second step in the purchase procedures.
- 'delivered': the number of orders confirmed and delivered to their buyers. However, here only the e-buyers who purchased the products are counted.

The data resulted from the company activity between December 2021 and February 26, 2022. Figure (4) reveals that the orders hit 2194, meaning the number of interested customers in a product (a set of products) provided by company X. It indicates the considerable e-marketing efforts done by the company. A simple calculation of the total marketing costs based on this data knowing that m = 300 da, orders = 2194, the total marketing costs = 657600 da. The marketing costs are enormous for a relatively small company containing only four employees, including the business owner. Also, we notice that the number of confirmed orders is almost half of the orders, which means that phone calls were made with all those who ordered and only 1292 confirmed their orders, the phone calls costs are unknown, but a loss is observed. The figure also shows that only 861 purchases have passed all the purchase procedures, and the customers paid for their products. We can calculate the shipping costs covered by the company, which are uncompleted by a purchase. Where customers confirmed and did not finish their purchases: 'confirmed' – 'delivered' = 431

uncompleted purchases, knowing that r: the shipping cost = 350 da, loss in shipping = 150850 da, which is an important number. In this example, data on the company's profits are unavailable to compare the profits and the costs and to capture the real impact of these costs on the company's financial situation, which is out of our scope. However, we can see through these digits the impact of the e-customers behaviors on X and the costs driven by the different behaviors.

Fig. 4. E-customers in purchase: 'delivered', 'confirmed', and total.



We note that it is a case of information asymmetry (where the e-sellers are not informed about the e-buyers). Even if the e-buyers are appeared to be interested, take their time to know everything about the product, and confirm their orders, they may present another decision when the product is delivered to them. As we previously mentioned, they even cancel their orders at a later purchase stage. The example shows that the percentage of e-buyers who performed their purchases is only 39%, whereas e-buyers who confirmed their orders are only 58%. We consider that there are three types of e-customers. We refer here to the term e-customers as people who enter the platform and make an order of a product. The quality of the e-customer behaviors is different. However, the company has different risks in dealing with e-customers depending on the quality of the behaviors they have.

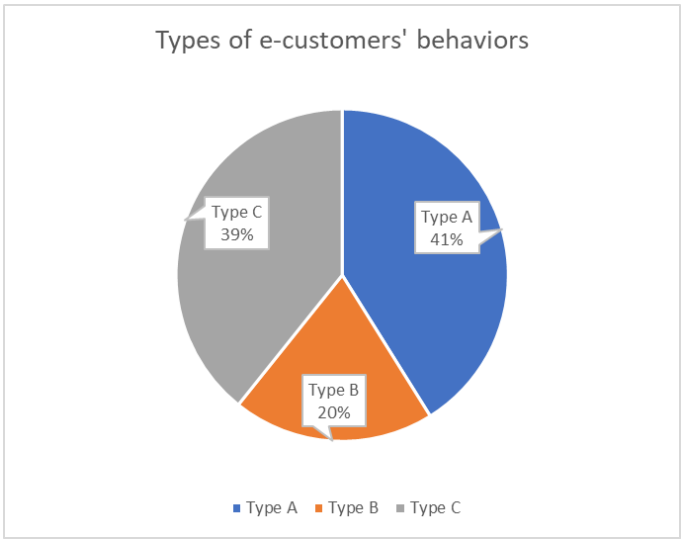
A: e-customers with the behavior of making an order and then not confirming it.

B: e-customers with the behavior of making an order, confirming it, and not finishing their purchases

C: e-customers with the behavior of finishing their purchases

According to the X case in the period mentioned, the proportion of individuals of type A s more remarkable than those of type B greater than those of type C, as shown in figure (5).

Fig. 5. Types of e-customers based on their behaviors in dealing with the local e-commerce example (type A, type B, type C)



In the beginning, we thought of using 'commitment' so people who complete their purchases are more committed, in contrast to those who cancel their purchase in step one or two. However, the word commitment is vague and may not be the most appropriate to describe this behavior, and we use the term 'behavior'.

The purchase procedures used in this type of e-commerce, notably the case presented, complicate adapting the lemons model to the local e-commerce market in Algeria. However, this problem is a massive concern to e-sellers as potential buyers introduce risks to the company even when they are not purchasing, which is the difference between the real market and the e-market. In local e-commerce, e-customers behavior is essential, and it influences companies not only the profit and financial situation but also their dealings with future e-customers. If a considerable number of e-customers ordered a specific product and the majority of them confirmed (type B), when the company delivers the products to e-customers, it moves the products from its stocks, leading to a lack of this product. Consequently, it cannot confirm orders of future e-customers who ordered this product till the return of the products if the e-customers are of type B. Many issues may appear in the presence of this variety of behaviors from the e-customers. Nevertheless, it is noted that local e-sellers cannot measure this behavior of their customers. They are likely to increase the prices of their products to cover the costs resulting from the unknown behaviors of e-buyers.

A proposed approach on dealing with information asymmetry in the case study. An approach we propose to deal with this issue, inspired by the Switzerland project to reduce information asymmetry in the used cars market (Zavolokina, Schlegel and Schwabe, 2021), is a platform for e-buyers and e-sellers, each e-buyer or e-seller has a unique account on this platform which can easily be controlled by requiring the ID number while registration. The concept of this platform is to help e-sellers recognize the e-buyer behavior before the second step and the third step in the purchase, using the rating for e-buyers and e-sellers as well. So, each e-seller will rate his dealing experience with a certain e-buyer (rate from 1 to 5), and the same for e-buyers can rate their experiences with e-sellers. A high rating for an e-buyer refers to his good behavior

in purchasing from local e-commerce. An e-seller with a high rating indicates its product's quality and service. It is proposed that ratings will help both sides, whether e-sellers or e-buyers, well inform about the other side in dealing, reduce information asymmetry, and save local e-commerce from inevitable losses without knowing their prospective e-buyers.

A solution based on the market of lemons and information theory. This solution applies a model proposed by Robert (Mamada, 2021). According to Spence's signaling theory, signals in job markets can solve adverse selection. Robert suggests that these signals are vital in solving adverse selection. At the same time, he also incorporates the costs of knowing these signals in sequential Bayesian games. He calculated these costs using Shannon information and the entropy of signals. We apply a considerable part of this model in the solution we propose. We first measure the entropy. Here, it refers to the uncertainty, the conditional entropy, and the mutual information. In our exercise, we extracted the games as in the paper Robert, but since our case is different, regarding our modest investigation and the scope of this paper, the results (if they are true) are not analyzed.

Table 3. The joint probability distribution

Age Group (G)\ Behavior (V)	A	B	C	$P(G_i)$
G_1	p	q	r	$p+q+r$
G_2	s	t	u	$s+t+u$
$P(V)$	$p+s$	$q+t$	$r+u$	

In our example, many indices can serve to help the e-seller inspect whether the e-buyer has good behavior or not, means of type C or not, such as the 'city', 'age', 'degree of interest' ('commitment'), 'educational level', etc. For simplicity purposes, we take as an index the age group, assuming that people under 30 years old have good behavior since they are more familiar with the purchase procedures than people older than 30 years old.

Table (3) refers to the joint probability distribution, this table can be extracted based on historic data with 2194 observations, so p, q, r, s, t, u can be defined easily. A, B, and C indicate 'behavior of type A', 'behavior of type B', and 'behavior of type C' respectively. $G_1(G_2)$ denotes 'the first age group (≤ 30)' ('the second age group (> 30)') of the e-buyer. According to the data we dispose of, we assume $p > q > r$ and $s > t > u$, and we assume that $u < s, t, q, p$ and $r < q, s, t, p$.

At this point, implementing the model mentioned earlier, the Shannon information is calculated for each bit. For example, the Shannon information stands for G_1 in table number (3), which designs the information amount telling us that the index in 'the first age group' $= -log_2(p+q+r)$. For $G = \{G_1, G_2\}$, $V = \{V_1, V_2\}$, the Shannon information standing for G_2 (the information amount telling us that index is 'the second age group') $= -log_2(s+t+u)$, Shannon information standing for A (the amount of information telling us that the buyer's behavior is of type (A))$= -log_2(p+s)$.

As described in Robert's research, Entropy in this case is used to measure the uncertainty, then the entropy (uncertainty) reflects the index's group. Consequently, uncertainty vanishes if the index type is known (age group). The entropy about the information of the index (G) is $H(G) = -(p+q+r)log_2(p+q+r) - (s+t+r)log_2(s+t+r)$, the entropy about the behavior of the e-buyer is: $H(V) = -(p+s)log_2(p+s) - (q+t)log_2(q+t) - (r+$

$u)log_2(r+u)$. Since V and G are both discrete variables, their joint entropy corresponds to their pairing entropy (V,G), therefore $H(V,G) = -plog_2(p) - qlog_2(q) - rlog_2(r) - slog_2(s) - tlog_2(t) - ulog_2(u)$, the conditional entropy indicate uncertainty average, not attributed to G, in V. The uncertainty (entropy) of V knowing that G_i has happened is $H(V|G_i) = -\sum_{i,j=1}^{2,3} P(G_i)P(V_j|G_i)log_2(P(V_j|G_i))$. The uncertainty of V after G is $H(V|G) = -\sum_{i,j=1}^{2,3} P(G_i)P(V_j|G_i)log_2(P(V_j|G_i))$. Therefore:

$$H(V|G) = -plog_2\left(\frac{p}{p+q+r}\right) - qlog_2\left(\frac{q}{p+q+r}\right) - rlog_2\left(\frac{r}{p+q+r}\right) - slog_2\left(\frac{s}{s+t+u}\right) - tlog_2\left(\frac{t}{s+t+u}\right) - ulog_2(\frac{u}{s+t+u})$$

The mutual information (I) between V and G $I(V,G)$ means the amount of uncertainty removed by G about V and vice versa. $I(V,G) = H(V) - H(V|G)$:

$$H(V) = -(p+s)log_2(p+s) - (q+t)log_2(q+t) - (r+u)log_2(r+u)$$

$$I(V,G) = -(p+s)log_2(p+s) - (q+t)log_2(q+t) - (r+u)log_2(r+u) + plog_2(p) + qlog_2(q) + rlog_2(r) - (p+q+r)log_2(p+q+r) + tlog_2(t) + slog_2(s) + ulog_2(u) - (t+s+u)log_2(t+s+u)$$

Based on this, according to the models suggested by Robert, he suggests two games one in which the buyer inspects the index quality, and in the second, the sellers tell the buyer about the index quality. In our case, the extracted games may not be reasonable according to the model of Robert. We proposed that e-sellers require as many indexes as possible from the e-buyers in the first step of purchase 'ordering'. These indexes can identify if e-buyers are from class C, which corresponds to buyers with good behavior and completed purchases. The most discriminative indexes can be selected using the mutual information scores. The higher the mutual information value, the more influential the variable is in recognizing the behavior of the e-buyer. Even a classification model can be employed to classify e-buyers behavior. In this case, we can group types A and B in one class, calling D referring to bad e-buyers' behavior and C referring to buyers with good behavior in the purchase. As a result, it will be a binary classification task. A simple decision tree model, a well-known data mining algorithm for classification based on the entropy criteria, can predict the class of the e-buyers using the selected indexes. As we mentioned before, 'educational level', 'age', 'city', 'previous purchase from e-commerce' might be informative indexes to tell about the e-buyer behavior in the purchase, reducing the uncertainty about e-buyers in the purchase. Consequently, local e-commerce reduces their risks in dealing with e-buyers. At least it will know the behavior type of the e-buyer with a certain precision.

5 Conclusion

Asymmetry of information leads to adverse selection in almost all markets (used automobiles, tourism, pharmaceutics, e-commerce, etc.). As we have seen with our example of local e-commerce, the asymmetry of information emerged in the e-market. The lemons model should be adapted to these new contexts of information asymmetry. However, in our example, asymmetry of information happens in distinct ways, not only in the purchase, as the purchase procedures may differ from the one in the real market. As a result, solutions to lower the information

asymmetry inspired by the literature are suggested, like a decision tree model and the local e-commerce platform of e-buyers and e-sellers in Algeria. Further investigation can explore the use of data mining to lessen information asymmetry.

References

Akerlof, G. A. (1970) 'The market for "lemons": Quality uncertainty and the market mechanism', *Quarterly Journal of Economics*, 84(3), pp. 488–500. doi: 10.2307/1879431.

Baggio, R. and Baggio, J. A. (2011) 'Experiencing information asymmetries in tourism', in *4th Advances in tourism marketing conference (ATMC)*. Available at: https://digitalcommons.usu.edu/envs_facpub/1535 (Accessed: 28 February 2022).

Bond, E. (1982) 'A Direct Test of the "Lemons" Model: The Market for Used Pickup Trucks', *The American Economic Review*, 72(4), pp. 836–840. doi: 10.2307/1810022.

Hoffer, G. E. and Pratt, M. D. (1987) 'Used vehicles, lemons markets, and Used Car Rules: Some empirical evidence', *Journal of Consumer Policy*, 10(4), pp. 409–414. doi: 10.1007/BF00411482.

Kessler, A. S. (2001) 'Revisiting the lemons market', *International Economic Review*, 42(1), pp. 25–41. doi: 10.1111/1468-2354.00099.

Leland, H. E. (1979) 'Quacks, Lemons, and Licensing: A Theory of Minimum Quality Standards', *Journal of Political Economy*, 87(6), pp. 1328–1346. doi: 10.1086/260838.

Light, D. W. and Lexchin, J. R. (2021) 'Pharmaceuticals as a market for "lemons": Theory and practice', *Social Science & Medicine*, 268, p. 113368. doi: 10.1016/J.SOCSCIMED.2020.113368.

Lin, Y. (2012) 'Game theory research on the "lemon" problem of the e-commerce market', in *Advances in Intelligent Systems Research*. Atlantis Press, pp. 1513–1515. doi: 10.2991/iccia.2012.375.

Mamada, R. (2021) 'The Market for Lemons and Information Theory', *SSRN Electronic Journal*. doi: 10.2139/ssrn.3977876.

Ryoo, J. (1996) 'Lemons Models of Professional Labor Markets Reconsidered', *Eastern Economic Journal*, 22(3), pp. 355–363. Available at: https://www.jstor.org/stable/40325724 (Accessed: 26 February 2022).

Wilson, C. (1980) 'The Nature of Equilibrium in Markets with Adverse Selection', *The Bell Journal of Economics*, 11(1), p. 108. doi: 10.2307/3003403.

Zavolokina, L., Schlegel, M. and Schwabe, G. (2021) 'How can we reduce information asymmetries and enhance trust in "The Market for Lemons"?', *Information Systems and e-Business Management*, 19(3), pp. 883–908. doi: 10.1007/s10257-020-00466-4.