ISSN 1936-5330

kcFED Research Working Papers

An Experimental Analysis of Quality Misperception in Food Labels

Francisco Scott

October 2023
RWP 23-11
http://doi.org/10.18651/RWP2023-11

Federal Reserve Bank of Kansas City

An experimental analysis of quality misperception in food labels

Francisco Scott

Federal Reserve Bank of Kansas City

Abstract

The size and distribution of surplus in markets where credence quality attributes of food (e.g., organic, non-GMO) are conveyed through some informational mechanism (typically labels) crucially depend on 1) how information changes consumers' perception of quality and 2) producers' strategic choice of quality provision in response to changes in consumers' perception of quality. This paper examines the hypothesis that consumers' misperception of quality information can provide incentives to sellers to increase quality and offset the lower quality that exists in markets where firms imperfectly compete in quality and prices. Using previously derived theoretical predictions of a two-stage game in which firms sequentially choose qualities—which are misperceived by consumers—to then simultaneously choose prices, I conduct a laboratory experiment that emulates changes in consumers' perception of quality and examines their effects on producers' provision of quality and market surplus. My results indicate that total surplus increases mainly with overvaluation of the high-quality product, confirming theoretical predictions. But contrary to the theory, I find that low-quality sellers try to compete by raising their quality levels too much when low quality is overvalued, dampening their quality-adjusted prices. As a result, welfare approaches first-best only when the high-quality product produced by the market leader is overvalued. These results highlight the importance of examining the market structure when designing informational policies.

JEL classifications: C9, Q18

Keywords: quality competition, laboratory experiments, food labels

Information-based mechanisms identifying unobserved credence attributes (e.g., labels) have proliferated in the markets of food products. Their increasing popularity is based on the assumption that they help consumers making better choices by eliminating informational asymmetries (Roe and Sheldon, 2007; Bonroy and Constantatos, 2015; Lusk et al., 2018). Evidence provided by choice experiments, however, reveals that consumers often fail to understand the information conceived in labels (Kiesel and Villas-Boas, 2013; Lee et al., 2013; Liaukonyte et al., 2015; McFadden and Lusk, 2018) and often miss-perceive (over- or under-value) the true quality of labeled products. Economic intuition suggests that misperception can distort consumers' choices and change sellers' strategic responses (e.g. quality provision and prices charged), impacting the size and distribution of welfare in the market. Despite the importance of information transmission in food markets, there is a dearth of empirical studies examining sellers' strategic reactions to consumers' misperception. The lack of empirical studies is partly explained by the unobserved nature of consumers' misperception and the difficulty of finding data with observational variation to misperception (Dranove and Jin, 2010).

To circumvent limitations of data availability, this paper reports a laboratory experiment that examines how consumers' misperception about quality information alters sellers' strategic responses. This paper manipulates the intensity of consumer misperception in a laboratory experiment to test the direction and magnitude of the effects of misperception of quality on market outcomes (qualities and prices), and welfare outcomes (consumer surplus, profits, and total welfare). The rationale for doing so is as follows. Under perfect quality information and conditional on market structure, informational mechanisms (henceforth called labels, for simplicity) can increase market efficiency because consumers use labels to verify the quality of products and adjust premiums accordingly, avoiding rent-seeking behavior from sellers (Roe and Sheldon, 2007; Dulleck and Kerschbamer, 2006; Roe and Sheldon, 2007; Bonroy and Constantatos, 2015). Consumers' misperception (i.e., attribution of subjective quality valuation to a product because of limited capacity to capture label information) alters the marginal benefit of adopting different labels, allowing firms to strategically adopt a different quality label to capture rents.

Incentives to adopt different quality labels are relevant for several reasons. Firms do not capture all the social benefits of increasing quality in imperfectly competitive markets. Their decision to supply quality is based on the marginal consumers' willingness-to-pay (WTP) for quality, which is often different from the average consumers' WTP for quality (Spence, 1975; Mussa and Rosen, 1978; Champsaur and Rochet, 1989; Buehler and Schuett, 2014). The mismatch between average and marginal consumes' WTP implies that the level of quality offered by firms does not align with the quality level desired by society. But if firms are responsive to consumers' misperception of quality, the intensity of misperception can be high enough to incentivize firms to change quality provision and approximate a first-best welfare level. Conversely, informational-based policies (e.g., educational programs, nudges) that try to curb misperception can decrease the incentives for quality provision and lower total welfare as a result. This paper tests the impact of misperception on the size and distribution of welfare. It does so by analyzing the outcomes of an experiment that leverages on theoretical predictions in a model of vertical differentiation. Vertical differentiation is appropriate in this context because it matches empirical regularities in the use of quality labels, the most common information mechanism in in food markets (Bonroy and Constantatos, 2015).

Specifically, this paper's experimental design is based on the comparative statics from Scott and Sesmero (2022) that modify the canonical model of competition by vertical differentiation with costly quality provision to include the effects of consumers' misperception of quality. They consider a market where two single-product sellers— one serving the high-quality segment of the market and the other the low-quality segment— imperfectly compete on quality and prices. While their study only considers a duopoly market, they are able to include consumers' relative misperception between quality grades and generate unambiguous comparative statics from the intensities of consumers' misperception. This paper uses these comparative statics to design treatments in a experiment. The treatments emulate shocks of misperception by doing a pairwise comparison between the market and welfare outcomes of a market where misperception is absent (control) to the outcomes of four types of consumers' misperception (treatments). Treatments take the form of 1) relative overvaluation of the high-quality product; 2) relative undervaluation of the high-quality product; 3) relative overvaluation of the low-quality product; 4) relative undervaluation of the low-quality product.

The laboratory experiment shows that misperception can lead to higher average quality and total welfare, as suggested by the theory. The increase in average quality and welfare happens when consumer overvalue the high-quality product. Overvaluation of the high-quality product only increases quality provision by the high-quality seller, giving room for prices of the high-quality product to raise more than proportionally to quality. The increase of both quality and price allows for above normal growth of the high-quality seller's profit and, as a consequence, higher level of total welfare. But total welfare also increases in the absence of higher average quality. Under undervaluation of low-quality product, qualities and prices of both low- and high-quality products decrease. The price of the high-quality product decreases less than proportionally to quality while quality and price of the low-quality seller decrease proportionally, allowing the high-quality seller to increase its quality adjusted prices, its market share, its profits, and as a consequence, total welfare. The experiment also shows that at least one segment of consumers turns to be worse off regardless of the nature of the misperception shock.

A key corollary of our empirical findings is that changes in total welfare depends on the magnitude of the changes in the profit of the high-quality seller, which acts as the leader in setting quality standards. Theoretically, high-quality sellers accrue higher profits and higher market share in markets where firms differentiate vertically (Ronnen, 1991; Motta, 1993;

Aoki and Prusa, 1997; Lehmann Grube, 1997). This is what we find in our experiment and also an empirical reality in many markets of food products, such as cage-free egg markets in Brazil (e.g., Neves et al. 2021), private *vs.* national brands in beef in the Untied States, and cooperative for organic milk products (e.g., Su and Cook 2015). Because of the high-quality sellers' importance, shocks that alters quality and price strategy for the high-quality seller impact total welfare more.

This papers' results reveals that the firm supplying for the high-quality segment of the market is highly responsive to shocks in misperception of quality, significantly changing its quality and prices when misperception changes. By doing so, the the high-quality seller can significantly change its profit margin. This paper finds that the seller supplying for the low-quality segment over-commits to quality under all misperception shocks and cannot recoup margins with the subsequent price choices. This is particularly important for markets where low-quality sellers tend to obtain low margins and capture the fringes of the market.

My experimental framework is closely related to experiments examining quality commitment in markets of imperfect information. The experimental design of most of these studies either (1) exogenously varies the informational mechanism (e.g., Cason and Gangadharan 2002), (2) exogenously varies the quality signal that firms send to consumers (e.g., Henze et al. 2015), or (3) exogenously vary the number of sellers and the likelihood of fraudulent behavior in quality provision (e.g., Bonroy et al. 2019). Differently from (1), this paper examines an environment where firms already committed to a information certification program, such as labels, and do not rely on other types of quality persuasion as, for example, cheap-talk. The friction of the experiment presented here lies on the intensity of the mismatch between the information given by the label and the consumer perception of quality. This setting complements the literature related to (2). Papers in (2) examine imperfect information of quality affecting only a proportion of consumers that understand a quality signal. This paper, in contrast, analyzes the effects of the intensity of quality misperception. Finally, this paper does not allow for fraud in the supply side of the market, as in (3). Fraudulent behavior and free-riding are generally associated with collective action and reputation (Harbaugh et al., 2011; Bonroy et al., 2019) which, although important in our context, confound the treatment effects that the experi,ent presented here is after.

The results presented in this paper also contribute to a body of economic experiments examining markets of credence goods. The existent research have investigated how liability and verifiability alter the incentives to overcharge or mislead consumers about their necessities (e.g., Dulleck et al. 2011); how competition and incentives can alter the incentives to overcharge (e.g., Mimra et al. 2016); and other incentive problems related to the market of credence goods, as described in Kerschbamer and Sutter (2017). My study expands the literature on such market experiments, focusing on the intensity of consumers' misperception of quality. From a policy perspective, this paper reveals that correcting certain types of misperception can have a deleterious effect on efficiency and unpredictable consequences in the distribution of surplus.

I now turn to the theory that informs my experimental setting. After, I discuss the experimental procedures and the results.

1 Theory and experimental hypotheses

1.1 Equilibrium and comparative statics

We rely on the theoretical predictions of Scott and Sesmero (2022). This section reproduces the essential parts of their model and its intuition. The model considers a market where consumers differ in their taste for quality and are distributed uniformly along a continuum of willingness-to-pay (given by θ) for quality (given by v). The parameters of the uniform distribution are $[\underline{\theta}, \bar{\theta}]$. Quality is a credence attribute and, hence, unobservable to consumers.

Consumers rely on a credible, non-profit, third-party to certify quality grade v and the third-party uses a continuous grade program to certify quality. In the supply side, the model considers a single-product duopoly in which firms have access to the same technology, which consists of a constant marginal cost, normalized to zero, for simplicity. The firms offer products and use the third-party program to certify quality. The sunk cost of certification is represented by $C(\cdot)$. For this study, I resort to a quadratic cost structure, following previous papers (e.g. Motta 1993; Aoki and Prusa 1997; Buehler and Schuett 2014).

The model assumes an honest and non-strategic third-party, but allows for a relative misperception of certified grade quality v by consumers. For example, misperception can arise from imperfect disclosure or imperfect understanding of information of certified products, such that misperception creates a wedge between the actual quality offered by firms and the perceived quality by consumers, increasing the relative distances between actual and perceived distance between high- and low-quality products. Scott and Sesmero (2022) provides an extensive list of examples of food markets where this wedge appears. Most prominently these include how consumers overvalue the relative quality of high-products that contain multiple labels, how consumers undervalue high-quality of products that contain low-fat claim, how consumers overvalue low-quality products that make environmental friendly claims, and how consumers undervalue low-quality products that use cost saving technologies (to the producer) such as genetically modified inputs.

Since the model in Scott and Sesmero (2022) considers only two labeled products in the market, the quality of the product certified with the relatively higher quality grade is represented by v_h, and a relatively lower quality grade is represented by v_l, such that $v_h > v_l$. The model describes the two misperception parameters, k_h (misperception of the high-quality grade) and k_l (misperception of the low-quality grade). Perceived qualities are denoted by $k_h v_h$ and $k_l v_l$ for the high- and low-quality products, respectively. In the presence of overvaluation, the authors let $k_j > 1$ for product $j \in \{h, l\}$; in the presence of undervaluation $k_j < 1$; and in the absence of misperception, $k_j = 1$.

The model defines indirect utility of consumers that buy labeled quality grade j as $V_i(v_j, p_j, k_j) = \theta_i k_j v_j - p_j$, such that i index the consumer's position in the WTP distribution, θ_i is the consumer's valuation of quality, $k_j v_j$ is consumers' perceived quality of product $j \in \{h, l\}$, and p_j is the price of product j. The indirect utility of those consumers consuming the outside good is zero. This class of indirect utility is a modification of indirect utilities commonly found in the literature (e.g., Jean Tirole 1988; Lehmann Grube 1997; Bonroy and Constantatos 2015). Marginal consumers θ_{lh} (indifferent between low- and high quality), θ_{0l} (indifferent between outside good and low-quality), and aggregate demand functions (D_h, D_l) are derived by integrating the density of consumers over the boundaries defined by the marginal consumers (see Jean Tirole 1988).[1]

Conditional on the aggregate demands, firms compete in two stages. First, a quality-competition stage, in which firms choose quality. Then, a price-competition stage in which, conditional on quality, they compete in prices. The solution of the two-stage game is computed by backward induction in the usual way (Ronnen, 1991; Aoki and Prusa, 1997; Lehmann Grube, 1997).

Firms' profits are given by equations 1 and 2:

$$\pi_h = R_h(v_h, v_l, p_h, p_l; k_h, k_l) - C(v_h) = p_h D_h(v_h, v_l, p_h, p_l; k_h, k_l) - C(v_h) \qquad (1)$$

$$\pi_l = R_l(v_h, v_l, p_h, p_l; k_h, k_l) - C(v_l) = p_l D_h(v_h, v_l, p_h, p_l; k_h, k_l) - C(v_l), \qquad (2)$$

where $R_j(v_h, v_l, p_h, p_l; k_h,\ k_l)$ is revenue of the firm offering product j.

In this paper, price-competition is simultaneous, while quality-competition is sequential. The timing of the game reflects empirical situations in which a firm can commit to a specific quality before the other, but it cannot do the same with prices.

[1]The formal expression can be found in the Appendix.

Such timing is associated, for example, with markets where one firm can credibly commit to a quality standard before others. In food markets, this reflects the empirical realities of industries in which producing technologies are asset-specific, as the cage-free egg industry (e.g., EggIndustry 2019), or when large companies have better advertisement departments or reputation.

Scott and Sesmero (2022) shows that a sufficient condition for a global solution of the sequential programming described above consists in restricting misperpcetion k_h to $[0.75, 1.75]$ when $k_l = 1$, and k_l to $[0.5, 1.3]$ when $k_h = 1$. Under these parameters, the leader always assume the high-quality spectrum of quality, while the follower becomes the low-quality firm. The optimal quality solution of the game is represented by $\{v_h^*, v_l^*(v_h)\}$, which consists of the equilibrium quality chosen by the high-quality firm and the follower's best-response to the high-quality grade. Optimal prices are represented by $\{p_h^*(v_h^*, v_l^*), p_l^*(v_h^*, v_l^*)\}$. Total welfare is the summation of profit of the high-quality firm (π_h), profit of the low-quality firm (π_l), surplus of the segment of consumers purchasing the high-quality product (CS_h), and surplus of the segment of consumers purchasing the low-quality product (CS_l). Therefore, I define welfare as $W(v_h, v_l, p_h, p_l; k_h, k_l) \; = CS_h + CS_l + \pi_h + \pi_l$.

Notice that under misperception there is a divergence between the actual utility, defined as the one the consumer derives from the actual quality of the good, and the perceived utility, defined as the one the consumer derives from the perceived quality of the good. The authors follow the approach implemented in the literature (e.g. Glaeser and Ujhelyi 2010; Brécard 2014; Baksi et al. 2017) and evaluate consumer surplus based on the actual levels of quality provided, v_j, instead of the augmented perceived quality, $k_j v_j$. Formally, consumer surplus is defined as in equations 3 and 4.

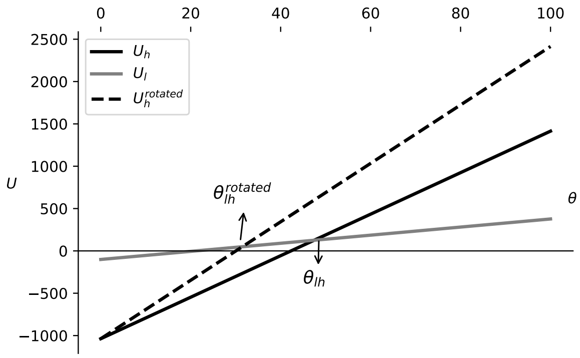
$$CS_h = \int_{\theta_{lh}}^{\bar{\theta}} \frac{\theta v_h - p_h}{\bar{\theta} - \underline{\theta}} d\theta. \tag{3}$$

$$CS_l = \int_{\theta_{0l}}^{\theta_{lh}} \frac{\theta v_l - p_l}{\bar{\theta} - \underline{\theta}} d\theta. \tag{4}$$

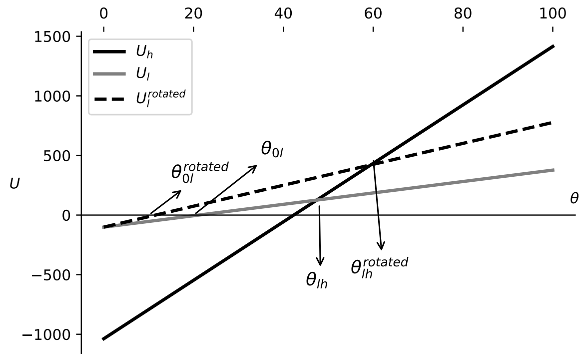
Armed with these definitions, I explore the effects of shocks in misperception (again, changes in k_h and k_l) to market outcomes (equilibrium qualities, prices, demanded quantity) and welfare (equilibrium profits and consumer surplus). Fundamentally, I consider changes in perception of relative quality by either allowing k_h to vary while holding $k_l = 1$ or the other way around. Figure 1 helps to describe the intuition behind these comparative statics, which are later tested. In all panels of the figure, the horizontal axis represents the consumer's WTP index, θ_i, and the vertical axis represents utility. The intercept of the utility curve represents equilibrium prices, and the curves' slope represents equilibrium quality.

Relative to the case where misperception is absent, the model predicts that a shock in which consumers start to overvaluation the high-quality product (an increase in k_h, holding $k_l = 1$) increases the perceived utility of high-quality consumers, as indicated by the counterclockwise rotation of its utility curve (figure 1a). Also, overvaluation of the high-quality grade strengthens the incentives for the high-quality firm to offer more quality. Thus, the high-quality consumer's utility further rotates left, expanding the market for the high-quality product, all else constant. This can be seen by the left shift of the marginal consumer θ_{lh} (figure 1c).

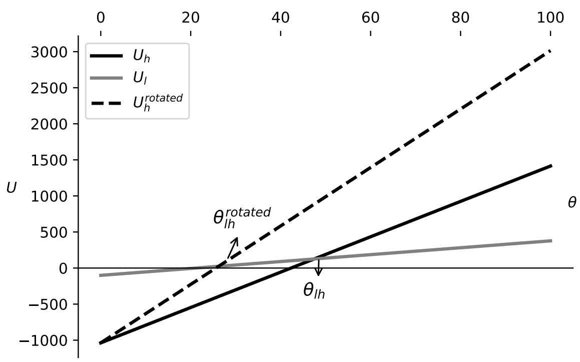
The increase in high-quality product allows the low-quality follower to capture part of the consumers with higher WTP by increasing the quality of the low-quality product, according to its best response function. Increase in quality implies a counterclockwise rotation of the low-quality consumers, as indicated by figure 1d. As a consequence of the higher quality products, both firms increase their prices (Figures 1e and 1f) until the marginal benefits (i.e., increase in markups) equates marginal costs (i.e., loss of market share and increase in sunk costs). It can be shown that at equilibrium firms are able to increase price more than they increase quality, which implies that quality-adjusted prices increase, and so do profits for both firms.



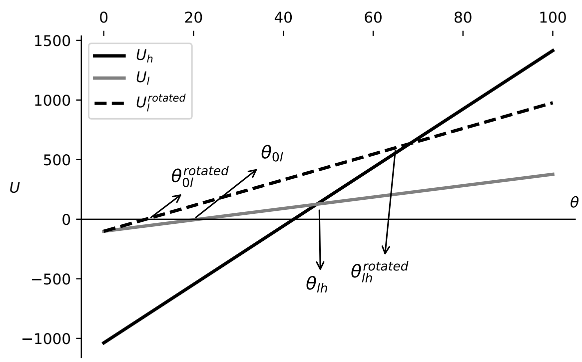
(a) Effect of overvaluation of high-quality product



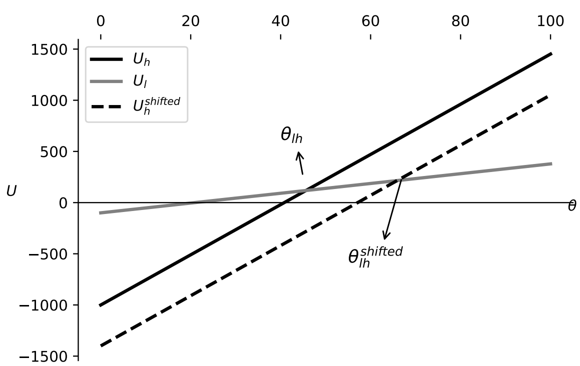
(b) Effect of overvaluation of low-quality product



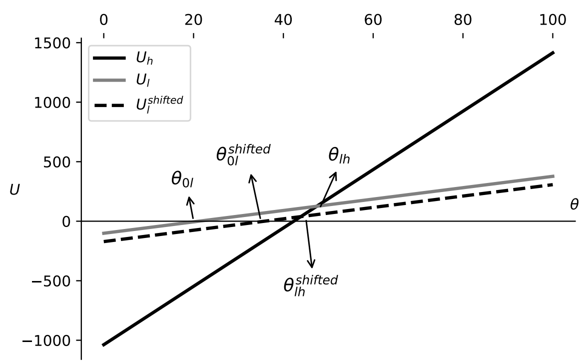
(c) Effect of quality increase in high-quality product



(d) Effect of quality increase in low-quality product



(e) Effect of price increase in high-quality product



(f) Effect of price increase in low-quality product

Figure 1: Effects of price and quality increase in the market of food labels. Each panel represents the effect of either prices or quality in marginal consumers, holding all else constant. Reproduced from Scott and Sesmero (2022) under different parameters.

Overvaluation of the low-quality grade (an increase in k_l, holding $k_h = 1$) rotates the low-quality consumer's utility counterclockwise (Figure 1b), strengthening the return to quality for the low-quality firm. By offering higher quality, the low-quality firm expands the market for its product, all else constant. However, to prevent losses in market share, the high-quality leader pre-emptively decreases its quality and prices. Additionally, the leader is able to decrease its fixed costs, as it only depends on the quality offered by the seller. In equilibrium, the model predicts that overvaluation of the low-quality allows the low-quality seller to increase quality, price, and its quality-adjusted price. This allows for higher profits for the low-quality seller. The high-quality firm decreases its quality and prices, in such magnitude that its quality-adjusted price decreases. As a result, its profits decrease.

The model predicts that consumer surplus decreases for the consumer segments incurring misperception. For example, high-(low-)quality consumers overpay for quality in markets where there exists overvaluation of high-(low-)quality labels. This implies a decrease in consumer surplus for this segment, according to equations 3 and 4. Putting all together, misperception is predicted to produce multiple forces impacting total welfare. While overvaluation (undervaluation) of high-quality increases (decreases) surplus for firms and low-quality consumers, it decreases (increases) surplus of the high-quality consumer segment. Scott and Sesmero (2022) show that these movements are related to two main variables: average quality supplied by sellers and total size of the market. The authors show that information-based policies that decrease average quality decrease welfare; but lacking increases in average quality, they also show that welfare can still increase if the size of the market expands enough to offset the deleterious effects of lower qualities to welfare.

Table 1 offers equilibrium values for all market and welfare variables based on the following parametrization: $\underline{\theta} = 0$, $\overline{\theta} = 100$ and over quadratic costs of the form $C(v_i) = \frac{v_i^2}{2}$, which numerically describes the intuition of the comparative statics. These values are also used in the experiment. For each case (i.e., over or under valuation of high- and low-quality products), the misperception parameters are set to one of the values on the boundaries the domain of k_j, holding $h_{-j} = 1$ for which a global solution is available. A formal derivation of these domains is in Scott and Sesmero (2022).

Table 1: Parameters and equilibrium solutions

	Parameter	Equilibrium	Surplus
Benchmark (BE)			
k_h	1	$v_h = 24.51$, $v_l = 4.78$	$\pi_h = 244.70$, $\pi_l = 15.15$
k_l	1	$p_h = 1037$, $p_l = 101$	$CS_h = 404.64$, $CS_l = 16.54$
		$p_h/v_h = 42.30$, $p_l/v_l = 21.12$	$TW = 681.01$
Overvaluation k_h (OH)			
k_h	1.5	$v_h = 37$, $v_l = 5.56$	$\pi_h = 628.66$, $\pi_l = 18.15$
k_l	1	$p_h = 2589$, $p_l = 128$	$CS_h = 80.57$, $CS_l = 19.16$
		$p_h/v_h = 69.67$, $p_l/v_l = 23.02$	$TW = 753.53$
Undervaluation k_h (UH)			
k_h	0.8	$v_h = 19$, $v_l = 4$	$\pi_h = 140.92$, $\pi_l = 13$
k_l	1	$p_h = 598$, $p_l = 80$	$CS_h = 425.62$, $CS_l = 13.78$
		$p_h/v_h = 31.47$, $p_l/v_l = 20$	$TW = 593.07$
Overvaluation k_l (OL)			
k_h	1	$v_h = 23.64$, $v_l = 5.12$	$\pi_h = 211.97$, $\pi_l = 21.47$
k_l	1.3	$p_h = 913$, $p_l = 128$	$CS_h = 438.29$, $CS_l = 10.49$
		$p_h/v_h = 38.63$, $p_l/v_l = 25.14$	$TW = 682.23$
Undervaluation k_l (UL)			
k_h	1	$v_h = 24.92$, $v_l = 3.64$	$\pi_h = 280.67$, $\pi_l = 7.41$
k_l	0.65	$p_h = 1155$, $p_l = 54$	$CS_h = 357.90$, $CS_l = 19.85$
		$p_h/v_h = 48.81$, $p_l/v_l = 22.82$	$TW = 665.83$
Common parameters to treatments			
$\bar{\theta}$	100		
$\underline{\theta}$	0		

v stands for quality, p stands for price, p/v stands for quality-adjusted prices, π stands for profit, $C.S.$ stands for consumer surplus and TW stands for total welfare. Subscript h refers to high-quality, and l to low-quality.

Next, I describe the hypotheses derived from these comparative statics.

1.2 Hypotheses

Based on the results of these comparative statics and the numerical values presented in table 1, I construct 5 hypotheses to be tested in an experiment. Hypotheses 1 and 2 refer to market outcomes: quality levels, and prices charged under different treatments of misperception of quality. Hypotheses 3-5 refer to welfare outcomes under the same treatments: firms' profits and consumer surplus of the high- and low-quality segments.

Hypothesis 1: Quality offered. In relation to the benchmark case in which misperception is absent (i.e., $k_h, k_l = 1$),

1. high- and low-quality increase (decrease) under overvaluation (undervaluation) of the high-quality product.
2. high-quality decreases (increases) and low-quality increases (decreases) under overvaluation (undervaluation) of the low-quality product.

Hypothesis 2: Prices. In relation to the benchmark case in which misperception is absent (i.e., $k_h, k_l = 1$),

1. high- and low-quality prices increase (decrease) under overvaluation (undervaluation) of the high-quality product.
2. The price of the high-quality product decreases (increases) and the price of the low-quality product increases (decreases) under overvaluation (undervaluation) of the low-quality product.

Hypothesis 3: Profits. In relation to the benchmark case in which misperception is absent (i.e., $k_h, k_l = 1$),

1. profits of the high- and low-quality firms increase (decrease) under overvaluation (undervaluation) of the high-quality product.
2. the profit of the high-quality firm decreases (increases) and the profit of the low-quality firm increases (decreases) under overvaluation (undervaluation) of the low-quality product.

Hypothesis 4: Consumer Surplus. In relation to the benchmark case in which misperception is absent (i.e., $k_h, k_l = 1$),

1. under overvaluation (undervaluation) of the high-quality product, the consumer surplus for the high- quality segment of the market decreases (increases), while consumer surplus of the low-quality segment of the market increases (decreases).
2. under overvaluation (undervaluation) of the low-quality product, the consumer surplus for the high- quality segment of the market increases (decreases), while consumer surplus of the low-quality segment of the market decreases (increases).

Hypothesis 5: Total Welfare. In relation to the benchmark case in which misperception is absent (i.e., $k_h, k_l = 1$), total welfare increases (decreases) under overvaluation (undervaluation) of high- and low-quality products.

These 5 hypotheses are tested in a laboratory experiment. The next section describes the experimental design.

2 Experimental Design

The experiment consist of a between-subjects design used to investigate the role of consumers' misperception on market and welfare outcomes. I compare the results of 4 treatments under different intensities of consumers' misperception to a benchmark case under the absence of misperception. Subjects take the role of firms, while the role of consumers is automated. Automated consumers allow for better causal identification of market and welfare outcomes because we eliminate possible behavioral confoundings that may arise from the demand side of the market.

The treatments take the form of (1) overvaluation of high-quality grade, (2) undervaluation of high-quality grade, (3) overvaluation of low-quality grade, and (4) undervaluation of low-quality grade. The misperception parameters are the boundaries of the domain of the misperpcetion parameters as described in table 1. Notice that the misperception (k_j,

$j \in \{h, l\}$) is a continuous variable. This implies that the experimenter can set the treatment level k_j anywhere in the interval where a global solution exists. To test the effects of undervaluation of the high-quality grade, k_h can be set anywhere in the interval $(0.75, 1]$. Likewise, the experimenter can choose any k_h between $(1, 1.75)$ to test outcomes under overvaluation of high-quality. To test undervaluation of low-quality, k_l can be set to any value in $[0.5, 1)$; overvaluation of low-quality needs k_l to be anywhere $(1, 1.3]$. To decide the appropriate levels of k_j, I follow List et al. (2011). The authors argue that, under continuous and monotonic linear treatment effects, the experimenter should set the treatment variable to extreme values, such that it maximizes the difference between treatment outcomes. Therefore, I set $k_h = 1.5$, $k_h = 0.8$, $k_l = 1.3$, and $k_l = 0.65$ for treatments (1), (2), (3), and (4), respectively. The exact experimental procedures can be discussed next with the parametrization and the numerical equilibria in hand (see table 1).

2.1 Procedures

The experimental sessions ocurred during September 2020 using oTree (Chen et al., 2016). Subjects are mainly undergraduates from a large university located in the United States. Student recruitment was managed via ORSEE (Greiner, 2015). I conducted 16 sessions, and targeted 8 students per session.[2] A session consisted of the following steps. First, the experimenter handled printed copies of the experimental instructions to subjects. The experimental instructions were read out loud. Second, subjects responded to a post-instruction quiz to check for their understanding of the rules of the game. Subjects were paid per every right question answered during the quiz. The experiment started after the post-instruction quiz. The experiment consisted of 2 phases: a training phase in which subjects played 4 rounds of the game (2 as leaders and 2 as followers) and an effective experiment that con-

[2]A session under benchmark had 10 students, and 5 others (under different treatments) had 6 students due to last minute cancellations.

sisted of 10 rounds. Each round is described according to figure 2. Each round consists of 3 periods: a period in which the leader makes its quality choice, a period in which the follower makes its quality choice, and a period in which leader and follower choose prices simultaneously. For each treatment, subjects face different incentives to provide quality and charge prices that are consistent with over- or undervaluation of high- or low-quality grades.

In practice, two subjects are randomly paired to play a round of the game. One of the players is randomly selected to play the leader (the high-quality seller), while the other plays the follower (the low-quality seller). The leader must select its own quality first; to facilitate the quality choice, the leader has access to a calculator that shows the revenue, costs, payoff, and the follower's payoff based on the qualities selected. To perform computations, the calculator requires the leader to guess the follower's quality and price choices, as well as a guess of its own price choices during the price period. The leader, thus, has 4 choices to make during its quality round: its quality, a guess for its own prices during the price period, and a guess for the follower's price and quality choices.

After the leader choices, the follower sees the leader's quality and chooses its own quality level. Again, a calculator with information about the follower's revenue, cost, payoff, and leader's payoff is available to facilitate the player's quality choice. The calculator uses the quality previously chosen by the leader to make its computations. To use the calculator, the follower has to choose its own quality, make a guess for its own price, as well as a guess the leader's price during the price period. Finally, during the price period, both players observe their quality choices and must choose prices for their product. Similarly to previous periods, a calculator is available. Players must choose their price and make a guess for the other player's price during the price period to use the calculator. After the price period, players observe their payoff, calculated via equation 1 and 2, and a new round starts. To make choices, subjects move a handle or type the quality and price values they wish to choose.[3]

[3]The appendix shows screenshots of how subjects select prices and quality.

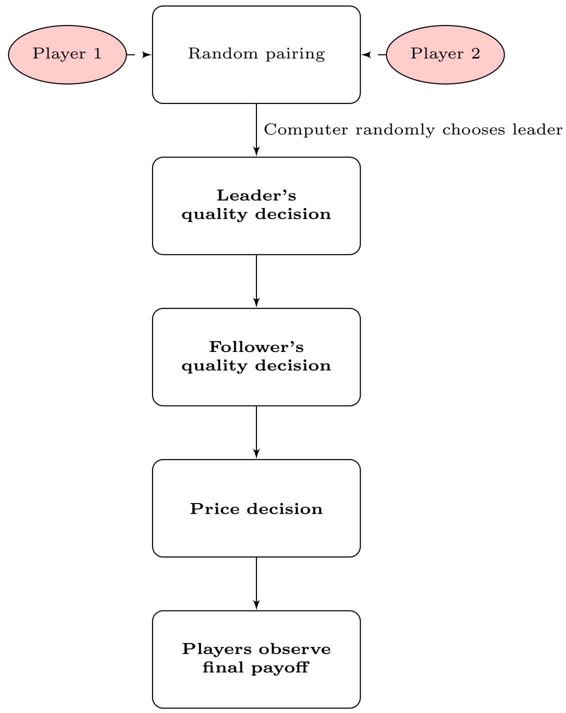


Figure 2: Description of a round

The decision space for high quality was set to $[16, 50]$, while the decision space for the low quality was set to $[2, 15]$. This decision is informed by the conditions by which the comparative statics were derived, which requires $v_h > v_l$. The decision space for price is set to $[590, 2700]$ for the high-quality product, and $[50, 150]$ for the low-quality product. Again, we restrict the decision space for values $p_h > p_l$. The decision spaces track the theoretical results and are chosen to minimize out-of-the-path equilibria that may arise from behavioral aspects of the game.

The experiment uses a payment schedule based on a random selection of rounds to be paid within a session. Out of the 10 effective rounds of each experimental session, the experiment randomly selects 4 to be paid. During the experiment, payoff values are named points, such that points are converted to U.S. dollars by a conversion rate. Subjects' average payment during the sessions, including the \$5 show-up fee, was \$16.92 for a 1-hour session.

The observed outcomes of the experiment are discussed next.

3 Results

3.1 Market outcomes

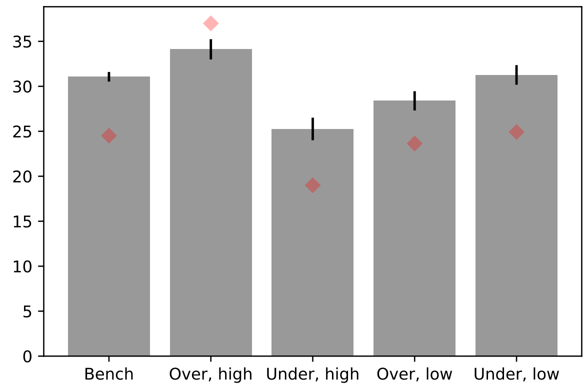
I discuss market outcomes first. Figure 3 summarizes the experimental results by looking at the means of different market outcomes.[4] Hypothesis 1 predicts that quality levels offered by sellers increase under overvaluation of the high-quality product; hypothesis 2 predicts the same for prices. Undervaluation predicts the opposite effects.

Notice that both the quality offered and the price charged by high-quality sellers under the benchmark experiment are higher than what the theory predicts (Figure 3a and 3b). Previous papers report a similar pattern for experiments that studies quality competition (e.g., Henze et al. 2015). The data confirms that the high-quality seller substantially increases quality and prices under overvaluation of the high-quality product relative to theory. The opposite is true under most of the other treatments; qualities and prices for the high-quality seller tend to be lower than the theory predicts.

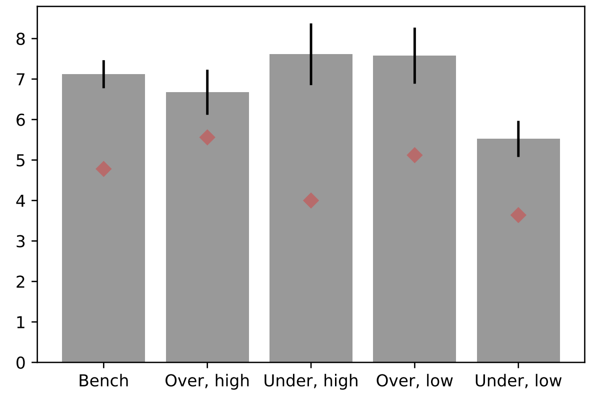
Interestingly, observed quality-adjusted prices (p_h/v_h) for the high-quality seller were much closer to the theoretical predictions (Figure 3e) than quality or prices taken separately. Recall that profits are determined by the size of the markup and also by the seller's market share, which is positively affected by qualities and negatively affected by prices. The quality-adjusted price shows how well sellers were able to balance the opposing forces enacted by changes in markup and market share. The closer quality-adjusted prices are to the theoretical predictions, the closer to the optimal profit sellers become, even if quality and prices are

[4] While the figures in this paper show means and confidence intervals, I also computed significance levels for the difference in means between benchmark and treatment, for all treatments. During these calculations, I corrected p-values for family-wise error rate (FWER) as described in List et al. (2019). Qualitatively, results are largely the same as presented here and, thus, not reported.

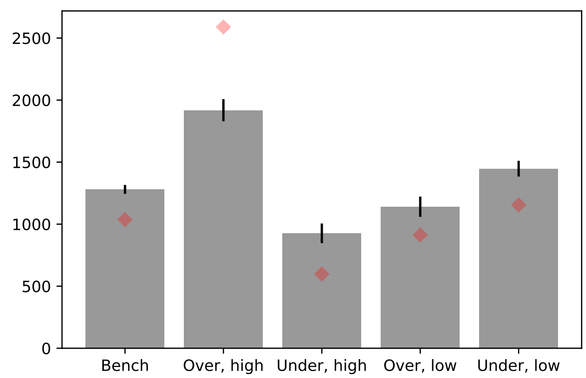
individually away from predictions.



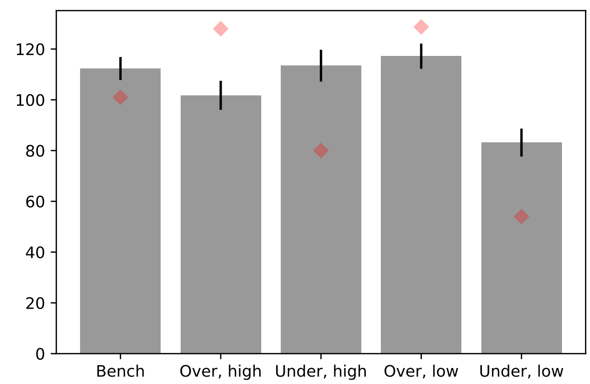
(a) Quality of high-quality product



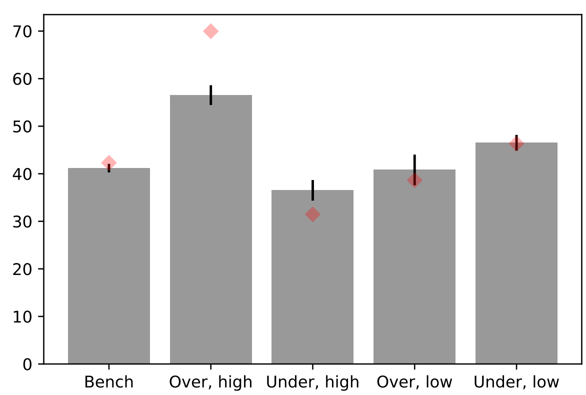
(b) Quality of low-quality product



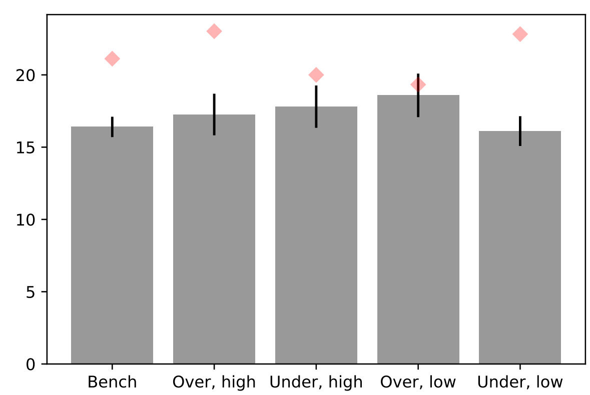
(c) Price of high-quality product



(d) Price of low-quality product



(e) Quality-adjusted prices of the high-quality product



(f) Quality-adjusted prices of the low-quality product

Figure 3: Market outcomes under different treatments. The height of the columns represent average observed outcomes, the red marks represent theoretical equilibria, and the bars are the 95% C.I.

While high-quality sellers' quality-adjusted prices are close to the theoretical predictions, low-quality sellers' quality-adjusted prices are consistently below (figure 3f). Particularly, low-quality sellers tend to offer qualities above what theory predicts (figure 3b). Prices charged by the low-quality seller are below what theory predicts for overvaluation of both products and above what theory predicts for undervaluation of both high- and low-quality products (figure 3d). The means for low-quality seller's quality, prices, and quality-adjusted prices are not different from the benchmark, except for prices under low-quality undervaluation.[5]

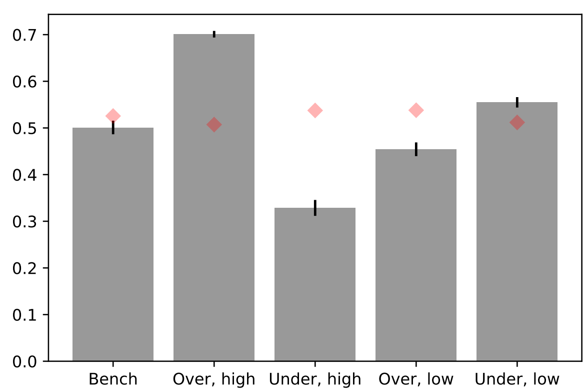
I explore two possible explanations for the lack of difference in market outcomes for the low-quality product. The first relates to the difficulty of the game for low-quality seller. This stems from low-quality being the follower in this market. Positive profits for the low-quality seller are possible in an smaller interval of the quality space if compared to the high-quality seller. More strain put on the decision of subjects playing the low-quality seller can imply a higher dispersion in quality decision. Notice also that the interval of optimal choices of low quality quality across treatments varies from 3.68 to 5.56, according to 1.

Together, a large dispersion of quality decisions allied with a small interval of optimal choices increase the likelihood of finding a null effect between treatments. This is true even with enough ex-ante statistical power to detect differences in means (by which the realized variance during the experiment is unknown). I do not find a large dispersion of quality choices around the mean across treatments. Without a large dispersion, the hypothesis that the game was too difficult for the low-quality seller is less likely. Nevertheless, I also conduct a battery of experimental sessions under which quality and price choices are set to discrete values to decrease the cognitive burden. The results, in the appendix, qualitatively agree with the results presented in the main results.

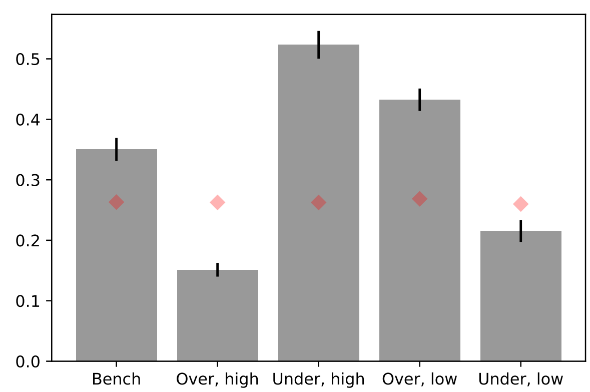
[5]Results described in figure 3 are further confirmed by a set of linear regressions models models used compare the different treatments with the observed qualities and prices obtained in the benchmark (no misperception of qualities). See the appendix for details.

A second explanation is behavioral. Higher-than-expected low quality in the quality round reveals a failure of backward induction. Conditional on a guess for prices, subjects playing low-quality seller decided for higher-than-expected qualities, probably because of higher quality positively impacted market share (as explained in Figure 1). The hypothesis becomes stronger when one examine the seller's price choice, for which the mean is mostly around 100.

The values of quality-adjusted-prices of the high-quality seller allied with a lack of best response behavior of the low-quality seller leads to its quality-adjusted-prices to be below the predicted optimal values. Again, quality-adjusted-prices away from the predicted values produce either too low market share for the low-quality seller (e.g., as seen in overvaluation of the high-quality and undervaluation of the low-quality in figure 4b), or too high market share but low mark up (as in undervaluation of the high-quality and overvaluation of the low-quality in figure 4b).



(a) Market share of high-quality product



(b) Market hare of of low-quality product

Figure 4: Market share under different treatments. The height of the columns represent average observed outcomes, the red marks represent theoretical equilibria, and the bars are the 95% C.I.

Products that need to be certified to reveal a hidden attribute but are relatively lower quality can be found in food and agriculture markets, particularly in the markets of produce or products that do not meet all the production quality production standards for expected by consumers (e.g., meat and eggs). The experiment shows that the interplay between margins and markups can have a serious impact in the profit of these low-quality sellers.

In summary, results presented here suggest that treatment effects (misperception of the high- or low-quality product) on low-quality choices (quality and price) are only strong enough when misperception direct affects the low-quality product. But misperception of either high- and low-quality products produce strong enough incentives to alter the high-quality seller's choices. Expected treatment effects for high-quality sellers aligned with unexpected magnitudes of treatment effects for low-quality sellers are likely to produce unexpected welfare results. I turn to welfare outcomes next.

3.2 Welfare outcomes

Hypothesis 3 predicts that overvaluation increases the profits of firms producing the misperceived product. Undervaluation is predicted to have the opposite effect on profitability. Hypothesis 4 predicts decreases in the surplus of the consumer segment that suffers from overvaluation of quality. For example, high-quality consumers are predicted to be worse off as they overvalue the high-quality product because they would be mistakenly overpaying for each unit of quality acquired. Hypothesis 5 states that overvaluation increases total welfare in the market, as overvaluation provides enough incentives to overcome underprovision of quality.

The previous sections showed that market outcomes under one of the treatments are, on average, different from the benchmark in the expected direction. However, welfare outcomes are a result not only of the direction of the treatment effect, but also of its magnitude. Tables 2 and 3 show linear regression models that compare the different treatments with the observed welfare outcomes obtained in the benchmark (no misperception of qualities). For each round of the game I calculate the welfare outcomes: the profit for high-quality seller, profit for the low-quality sellers, the consumer surplus for the high-quality seller segment, the consumer surplus for the low-quality segment, and total welfare. I model these market outcomes as the dependent variable in the model. The independent variables include a dummy taking the value of one if the observation belongs to treatment and zero otherwise, a conditional mean (intercept), and a time trend indicating the round during which subjects were making the choice, and random effects at the session level. Subjects' demographic characteristics were added to balance the samples, but suppressed from the table as they offer no particular insight. The sign of the treatment dummy ("Treat") indicates the direction of the misperception treatment over the dependent variable, and the magnitude indicates the impact of quality supply or price charged under each treatment.

The treatment effect for high-quality profits has the expected sign for all treatments. Compared to theoretical results, it underestimates the magnitude of the effect under overvaluation of high-quality, and it overestimates the magnitude under the other treatments. On the other hand, treatment effects for the low-quality profits have the expected sign for over- and undervaluation of the low-quality product, but the wrong sign for over- and undervaluation of the high-quality product. These results are a direct outcome of (1) the lower-than-expected low-quality prices charged under overvaluation of high-quality, and (2) higher-than-expected low-quality prices for undervaluation of high-quality associated with higher-than-expected observed qualities under all treatments. The failure to adjust for incentives provided by misperception often led some low-quality sellers to obtain negative payoffs during rounds.

I find no significant treatment effect for consumer surplus for the high-quality segment, except for undervaluation of the high-quality product. These results, which are contrary to the signs predicted by theory, arises from the much lower-than-expected prices charged by the high-quality seller. As a result, consumers were able to enjoy higher quality (Figure 3a), at a

Table 2: Treatment effects of welfare measures under misperception of the high-quality product

	Overvaluation of high-quality				
	Profit, high	C.S., high	Profit, low	C.S., low	Welfare
Const	162.6***	417.5***	9.277	31.12	679.2***
	(9.06)	(6.73)	(1.21)	(1.09)	(15.11)
Treat	252.8***	6.069	-19.68***	-47.23***	189.3***
	(33.79)	(0.47)	(-12.24)	(-7.81)	(37.42)
Round	-0.882	-2.017	0.174	3.070	0.258
	(-0.74)	(-0.54)	(0.47)	(0.86)	(0.20)
N	240	240	240	240	240

	Undervaluation of high-quality				
Const	118.6***	411.7***	5.843	42.09	655.2***
	(3.92)	(6.33)	(1.21)	(1.38)	(13.48)
Treat	-218.6***	-87.77***	17.29***	55.37***	-237.0***
	(-26.20)	(-5.72)	(11.09)	(6.67)	(-28.84)
Round	-0.671	-1.395	0.143	2.653	0.758
	(-0.54)	(-0.33)	(0.39)	(0.74)	(0.40)
N	230	230	230	230	230

*** prob. < 0.01, ** prob. < 0.05. Linear models estimated using hierarchical random effects (at the session level). Standard errors clustered at the session level. Subjects' demographic characteristics included in all specifications.

Table 3: Treatment effects of welfare measures under misperception of the low-quality product

	Overvaluation of low-quality				
	Profit, high	C.S., high	Profit, low	C.S., low	Welfare
Const	184.3***	471.3***	5.250	14.61	717.1***
	(5.76)	(5.99)	(0.83)	(0.39)	(11.42)
Treat	-66.83***	-7.427	8.938***	7.045	-58.14***
	(-6.03)	(-0.65)	(6.66)	(0.91)	(-6.16)
Round	-0.923	-4.247	0.460	3.745	-1.111
	(-0.70)	(-1.11)	(1.79)	(1.08)	(-0.59)
N	250	250	250	250	250

	Undervaluation of low-quality				
Const	149.4***	458.9***	6.964	30.21	699.7***
	(8.89)	(7.60)	(1.02)	(1.05)	(13.98)
Treat	61.35***	23.43	-14.76***	-28.13**	45.31***
	(14.17)	(1.74)	(-8.78)	(-2.91)	(5.93)
Round	-0.206	-1.398	0.0238	2.789	1.030
	(-0.19)	(-0.35)	(0.06)	(0.77)	(0.66)
N	230	230	230	230	230

*** prob. < 0.01, ** prob. < 0.05. Linear models estimated using hierarchical random effects (at the session level). Standard errors clustered at the session level. Subjects' demographic characteristics included in all specifications.

relative lower price (Figure 3c). The combination of qualities and prices for undervaluation of high-quality left the share of the market consuming the low-quality product way below what was predicted (Figure 4b). In combination with a higher-than-expected surplus under the benchmark, this explains the negative treatment effect on consumer surplus of the low-quality segment. A low market share for the low-quality product was also observed for the treatment effect under low-quality undervaluation, such that the treatment effect was significantly lower than the benchmark. The opposite happens under undervaluation of high-quality. Under this treatment, the share of consumers buying the low-quality product was above what theory predicts, resulting in a significant and positive treatment effect, as seen in Table 3.

Finally, total welfare is significantly higher when consumer overvalues high-quality, and significantly lower when consumers undervalue high-quality products. This is in line with the theory. But contrary to predictions, welfare decreases under overvaluation of low-quality, driven by the large decrease in high-quality profits. This large drop in profits is not compensated by any significant increase in high-quality consumer surplus, as predicted, leading to an overall drop in total welfare. The exact opposite happens under undervaluation of low-quality: the sharp increase in high-quality profits is large enough to offset the decrease in low-quality profit and consumer surplus.

In sum, much like the market outcomes, the welfare outcomes for the high-quality seller support the theoretical predictions. But low-quality profits had the expected sign of the treatment effects only under over- and undervaluation of low-quality. Measures of consumer surplus are largely a function of market shares and, because of that, much more sensitive to the magnitude of quality and price choices. Because of that, consumer surplus measures do not track theoretical predictions well. Surplus outcomes are direct corollaries of quality and price choices by the sellers.

4 Conclusion and policy implications

There is little empirical evidence about how sellers' decisions vary when consumers misperceive quality, particularly in food markets. As a consequence, researchers still do not fully grasp the efficiency and distributional effects of misperception. This makes policy that tries to curb misperception unpredictable from an efficiency and distributional point of view. The challenge for empirical studies lies in the fact that misperception is the difference between a consumer's perceived quality of a product, and the quality the consumer would perceive had they had full understanding of the product's credence attributes. This measure is not readily observable by the researcher, which limits identification strategies. To circumvent this limitation, I report results from a laboratory experiment that leverages on predictions about consumers' misperception on welfare and distribution under an empirical prevalent market structure, i.e. oligopoly markets in which sellers commit to quality of a product (via certification and labels, for example) and compete in prices.

Using different misperception intensities for different products (high- or low-quality), I tested the theoretical predictions of Scott and Sesmero (2022). I summarized these predictions under 5 hypotheses that describe how market and welfare outcomes under different intensities of misperception vary in comparison to when misperception is absent. Table 4 summarizes the treatment effects obtained from the experiment. All market and welfare effects for the high-quality seller are aligned with the theory, which implies that the high-quality seller tends to offer more (less) quality and charge higher (lower) prices under overvaluation (undervaluation) of high-quality products. Also, the high-quality seller tends to decrease (increase) quality and price to preserve (expand) market share under overvaluation (undervaluation) of the low-quality product. The experiment found no significant effects on market outcomes for the low-quality seller in most of the treatments. The unresponsiveness of the low-quality seller impacts distributional outcomes such that most of the theoretical predictions for the low-quality seller are either null (lack of significant treatment effect), or with the reversed signed as predicted by theory.

Table 4: Theoretical prediction vs. observed outcomes of market and welfare results. The sign to the left of the dash shows the theoretical predictions while the the sign to the right of the dash shows estimated treatment effect during the experiment

Effect	**Overvaluation of high-quality**	**Undervaluation of high-quality**	**Overvaluation of low-quality**	**Undervaluation of low-quality**
Quality				
High	+/ +	-/-	-/-	+/Null
Low	+/Null	-/Null	+/Null	-/-
Price				
High	+/ +	-/-	-/-	+/+
Low	+/Null	-/Null	+/+	-/-
Profit				
High	+/ +	-/-	-/-	+/+
Low	+/-	-/+	+/+	-/-
Consumer Surplus				
High	-/Null	+/-	+/Null	-/Null
Low	+/-	-/+	-/Null	+/-
Welfare	+/+	-/-	+/-	-/+

This paper shows that welfare outcomes under different misperception treatments can substantially differ from theoretical predictions even if the majority of market outcomes (qualities and prices) agrees with the sign of theoretical comparative statics. The magnitudes of the changes in quality and price under different misperception intensities directly impact the distribution of surplus, rendering most of the theoretical predictions on consumer surplus either null or with treatment effects with the reversed sign. The unexpected effects on distribution show that policies that try to curb misperception need to be explicit about which segment of the market the policy is targeting, so that different parts of society can evaluate the policy's effects.

Additionally, and for the same reasons, policymakers need to be attentive of the magnitude of the changes in sellers' choices after a policy to curb misperception is implemented.

For example, overvaluation of the high-quality product seems to affect supply of quality of the high-quality only, with no serious consequences for surplus of the high-quality consumer segment, as was initially suggested by theory. Most of the interventions to curb misperception would lead to a combination of qualities and prices that would produce a null effect on high-quality consumer surplus. Undervaluation of either product does not seem to impact the surplus of the low-quality segment in the direction suggested by theory. However, the experiment suggests that high-quality seller benefits from overvaluation of high-quality product or undervaluation of the low-quality product, as predicted by theory.

The most deleterious effect for efficiency would be a correction of overvaluation of the high-quality product. However, correcting undervaluation of high-quality products would be advised, as high-quality profits and consumer surplus are lower under such conditions, impacting total welfare negatively. More generally, if policymakers are interested in the total size of welfare, focusing on policies that target high-quality sellers and the segment of high-quality consumers would be best (as long as their market share is also sizable), because these consumers' WTP for quality tend to be large and the size of those market segments is way above the size of surplus from low-quality segment of the market.

Finally, it is important to emphasize that the usual limitations of the insights of laboratory experiments to policy implementation apply to this study, particularly its possible limited external validity. The limited external validity renders from the response of subjects to the structure and parametrization of our game. But I believe that this paper sheds important light on features that can be explored by further field experiments and observational studies. First, the higher capacity of high-quality sellers to influence the size of total surplus in markets under consumer's misperception of quality; second, the necessity to consider a wide range of misperception treatments to assess distributional effects; and third, the necessity to focus on size and magnitude of the effects of a policy that tries to curb misperception. I believe in the urgency for incorporating all three in future research that deals with labels and information provision.

References

Aoki, R. and T. J. Prusa (1997). Sequential versus simultaneous choice with endogenous quality. *International Journal of Industrial Organization 15*(1), 103–121.

Baksi, S., P. Bose, and D. Xiang (2017). Credence goods, misleading labels, and quality differentiation. *Environmental and Resource Economics 68*(2), 377–396.

Bonroy, O. and C. Constantatos (2015). On the economics of labels: How their introduction affects the functioning of markets and the welfare of all participants. *American Journal of Agricultural Economics 97*(1), 239–259.

Bonroy, O., A. Garapin, S. F. Hamilton, and D. M. Souza Monteiro (2019). Free-riding on product quality in cooperatives: Lessons from an experiment. *American Journal of Agricultural Economics 101*(1), 89–108.

Brécard, D. (2014). Consumer confusion over the profusion of eco-labels: Lessons from a double differentiation model. *Resource and Energy Economics 37*, 64–84.

Buehler, B. and F. Schuett (2014). Certification and minimum quality standards when some consumers are uninformed. *European Economic Review 70*, 493–511.

Cason, T. N. and L. Gangadharan (2002). Environmental labeling and incomplete consumer information in laboratory markets. *Journal of Environmental Economics and Management 43*(1), 113–134.

Champsaur, B. Y. P. and J.-c. Rochet (1989). Multiproduct Duopolists. *Econometrica 57*(3), 533–557.

Chen, D. L., M. Schonger, and C. Wickens (2016). oTree-An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance 9*, 88–97.

Dranove, D. and G. Z. Jin (2010). Quality disclosure and certification: Theory and practice. *Journal of Economic Literature 48*(4), 935–963.

Dulleck, U. and R. Kerschbamer (2006). On Doctors, Mechanics, and Computer Specialists: The Economics of Credence Goods. *Journal of Economic Literature 44*(1), 5–42.

Dulleck, U., R. Kerschbamer, and M. Sutter (2011). The economics of credence goods: An experiment on the role of liability, verifiability, reputation, and competition. *American Economic Review 101*(2), 526–555.

EggIndustry (2019). Are cage-free egg sales hitting a celling?

Glaeser, E. L. and G. Ujhelyi (2010). Regulating misinformation. *Journal of Public Economics 94*(3-4), 247–257.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association 1*(1), 114–125.

Harbaugh, R., J. W. Maxwell, B. Roussillon, R. Harbaugh, J. W. Maxwell, and B. Roussillon (2011). Label confusion: The Groucho effect of uncertain standards. *Management Science* (June 2011).

Henze, B., F. Schuett, and J. P. Sluijs (2015). Transparency in markets for experience goods: Experimental evidence. *Economic Inquiry 53*(1), 640–659.

Jean Tirole (1988). *The theory of industrial organization.* The MIT Press.

Kerschbamer, R. and M. Sutter (2017). The economics of credence goods - A survey of recent lab and field experiments. *CESifo Economic Studies 63*(1), 1–23.

Kiesel, K. and S. B. Villas-Boas (2013). Can information costs affect consumer choice? Nutritional labels in a supermarket experiment. *International Journal of Industrial Organization 31*(2), 153–163.

Lee, W. c. J., M. Shimizu, K. M. Kniffin, and B. Wansink (2013). You taste what you see: Do organic labels bias taste perceptions? *Food Quality and Preference 29*(1), 33–39.

Lehmann Grube, U. (1997). Strategic choice of quality when quality is costly: The persistence of the high-quality advantage. *The RAND Journal of Economics 28*(2), 372.

Liaukonyte, J., N. A. Streletskaya, and H. M. Kaiser (2015). Noisy information signals and endogenous preferences for labeled attributes. *Journal of Agricultural and Resource Economics 40*(2), 179–202.

List, J. A., S. Sadoff, and M. Wagner (2011). So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design. *Experimental Economics 14*(4), 439–457.

List, J. A., A. M. Shaikh, and Y. Xu (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics 22*(4), 773–793.

Lusk, J. L., G. T. Tonsor, T. C. Schroeder, and D. J. Hayes (2018). Effect of government quality grade labels on consumer demand for pork chops in the short and long run. *Food Policy 77*(October 2017), 91–102.

McFadden, B. R. and J. L. Lusk (2018). Effects of the National Bioengineered Food Disclosure Standard: Willingness to pay for labels that communicate the presence or absence of genetic modification. *Applied Economic Perspectives and Policy 40*(2), 259–275.

Mimra, W., A. Rasch, and C. Waibel (2016). Price competition and reputation in credence goods markets: Experimental evidence. *Games and Economic Behavior 100*, 337–352.

Motta, M. (1993). Endogenous quality choice : Price vs . quantity competition. *The Journal of Industrial Economics 41*(2), 113–131.

Mussa, M. and S. Rosen (1978). Monopoly and product quality. *Journal of Economic Theory 18*(2), 301–317.

Neves, M. F., A. W. Gray, C. E. Lourenco, and F. A. Scott (2021). Mantiqueira: innovating and disrupting in the egg business. *International Food and Agribusiness Management Review 24*(1030-2021-211), 138–161.

Roe, B. and I. Sheldon (2007). Credence good labeling: The efficiency and distributional implications of several policy approaches. *American Journal of Agricultural Economics 89*(4), 1020–1033.

Ronnen, U. (1991). Minimum quality standards , fixed costs , and competition. *The RAND Journal of Economics 22*(4), 490–504.

Scott, F. and J. P. Sesmero (2022). Market and welfare effects of quality misperception in food labels. *American Journal of Agricultural Economics 104*(5), 1747–1769.

Spence, M. (1975). Monopoly , Quality , and Regulation. *The Bell Journal of Economics 6*(2), 417–429.

Su, Y. and M. L. Cook (2015). Price stability and economic sustainability-achievable goals? a case study of organic valley®. *American Journal of Agricultural Economics 97*(2), 635–651.

A Power analysis and payoff curves

Table 5: Optimal sample size based on a 80% power for a t-test

	Minimum Effect Size		**Optimal number of obs**	
	Quality	*Price*	*Quality*	*Price*
Bench x Overvalue, kh				
High-quality	13.0	1500	6	3
Low-quality	1.2	27	245	29
Bench x Undervalue, kh				
High-quality	11.0	120	8	233
Low-quality	1.2	20	250	36
Bench x Overvalue, kl				
High-quality	2.0	120	184	90
Low-quality	1.2	20	115	7
Bench x Undervalue, kl				
High-quality	2.0	120	164	212
Low-quality	1.2	20	152	23

Notes: I used a 50% proportion between control and treatments. Minimum effect sizes were calculated based on a pilot experiment conducted before the main experiment and they are not the exact difference between optimal theoretical quality and price choices. Variances used in the power calculations were also based on a pilot experiment. Data from the pilot was not included in the main results.

A.1 Payoff curves

B Experimental instructions

In addition to the $ 5.00 you are entitled for showing up to this experiment today, you may earn an extra amount of money to be paid in cash after the experiment, in private. The extra amount of money will depend partially on the decisions you make during this session, partially on the decision of others, and partially to chance.

In this experiment, the extra amount of money you can earn is called payoff. During the experiment, your payoff consists of points, which will be converted to U.S. dollars by the end of the experiment. Each 35 points you earn will be converted to 1 U.S. dollar.

You are not to talk out loud during the session. We ask you to put away any electronic device (phones, tablets, etc.) you may be carrying with you today. Those who do not comply will be asked to leave. We expect this experiment to last between 60-90 min.

Outline

1. We will go over the instructions.
2. You will answer a post-instruction quiz designed to see if you understood the instructions correctly. You will be paid $ 0.2 per right answer. You will answer the quiz in your computer.
3. You will start the experiment. The experiment is divided in two phases.
    (a) A training phase designed to get you accustomed to the game. In the training phase, you will only play against the computer. In the training phase, the computer choices are independent of your choices. You will not be paid in the training phase.

(b) Effective experiment. Your performance in the rounds of the effective experiment will determine your final earnings.

4. You will answer a brief survey about your demographic characteristics.

The experiment

In this experiment, there will be sellers and buyers interacting in the market place. You and everyone else in this room will be sellers. The buyers will be played by the computer. The experiment consists of **multiple trading rounds**. During each round you will interact with one, and only one, other seller in this room, chosen randomly by the computer. After every round, the computer will randomly re-group you with another person in the room.

As sellers, you have to make two choices. **First, you have to decide what is the quality of your product you will be offering to buyers. Later, after all sellers decided their quality, you will choose the price of your product**. During each trading round, you will have the role of **Seller A** or **Seller B**. Your role during the trading round is randomly chosen by the computer.

Quality choices are done sequentially: Seller A chooses the quality of its product first, and Seller B chooses quality second, after observing Seller A's quality choice. **After Seller B chooses its quality, Seller A and Seller B decide their prices simultaneously, without the knowledge of the other Seller's price**.

Your **payoff** during each trade round will be determined by 1) how many buyers choose to buy your product, 2) the price of your product, and 3) the cost of offering the quality level you chose. **The payoff is simply determined by how many buyers bought your product multiplied by the price (what we call sales revenue) minus the cost of offering quality.** Thus, the higher your revenue, the larger your payoff will be. Also, the larger your costs, the smaller your payoff will be.

Quality

If you are Seller A: You can choose any number between the minimum quality 16 and the maximum quality 50.

If you are Seller B: You can choose any number between the minimum quality 2 and the maximum quality 15.

Prices

If you are Seller A: You can choose any price between the minimum price 590 and the maximum price 2700.

If you are Seller B: You can choose any price between the minimum price 50 and the maximum price 150.

Number of buyers

The number of buyers buying your product depends on the quality you choose, the price you charge, the quality chosen by the other Seller, and the price the other Seller chooses to charge.

As a general rule: if you increase quality of your product, more buyers will buy from you; but the more the other Seller increases its quality, less buyers will buy your product. Similarly with price. As you increase the price of your product, less buyers buy your product, and as the other seller increases its prices, more buyers buy your product.

It is possible that at a given combination of qualities and prices, no buyers will want to buy from you. It is also possible that under a different combination of qualities and prices, all buyers will want to buy from you.

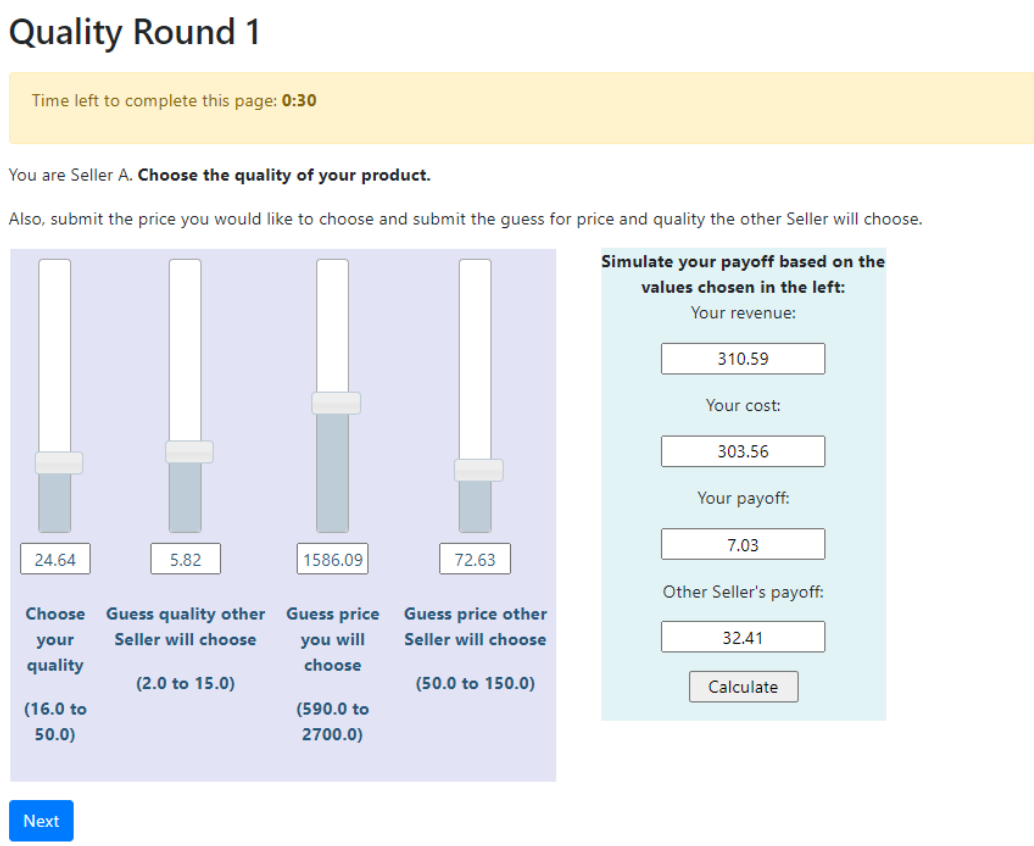
Costs

Your costs will depend on the quality you choose for your product. The higher the quality, the higher your costs will be. We will provide numerical examples in the next section.

Specific trading instructions

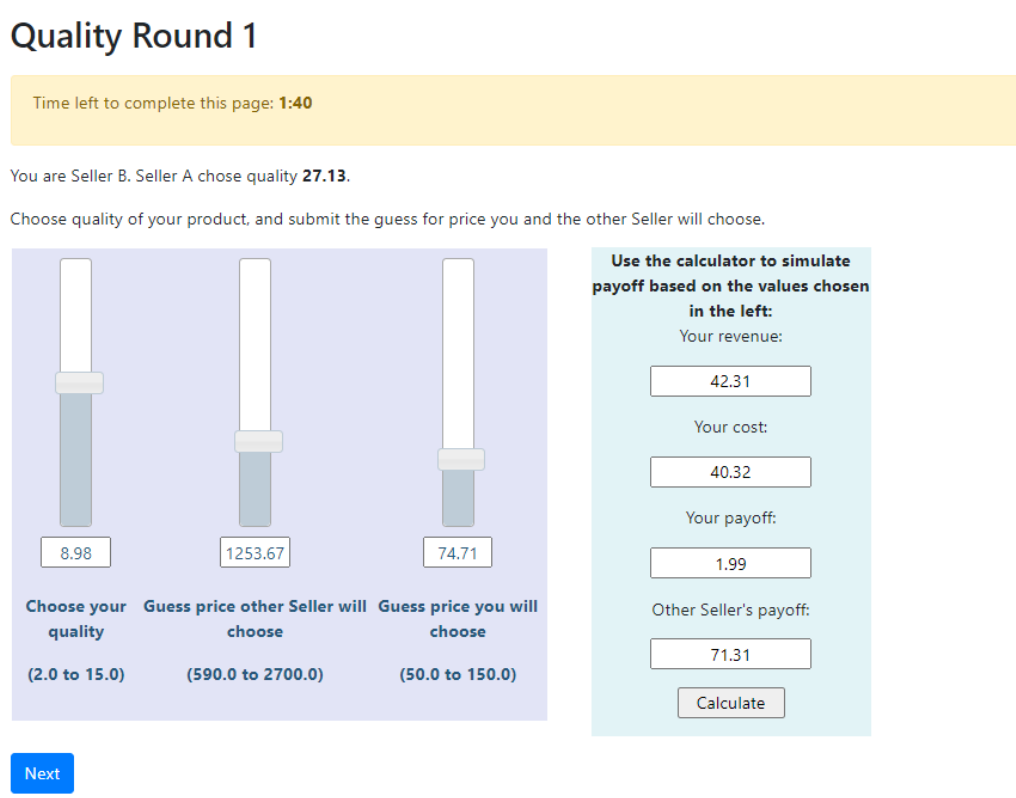
There will be 10 rounds. The sequence of choices in each round is:

- **Seller A quality choice:** Seller A will first choose the quality of its product. Additionally, we ask Seller A to a) make a guess for Seller's B quality choice ; b) choose the price they will choose if Seller B chooses the quality they guessed in (a); c) make a guess for the price Seller B will choose after it chooses the quality you guessed in (a). The screen in which Seller A makes the choices is depicted below. To make choices, click on one of the options.



The calculator indicates the payoff with the values Seller A chose by moving the handles. To use the calculator, make the choices, then click the button "Calculate". Click the button “Next” when ready to commit to choices. If you do not press “Next” before the time to complete this page, the computer will randomly make a selection for you. But there is no need to rush your decisions, as there is plenty of time to choose. While Seller A makes its choices, Seller B waits. Notice that Seller B will not be able to see Seller A’s choices while it waits.

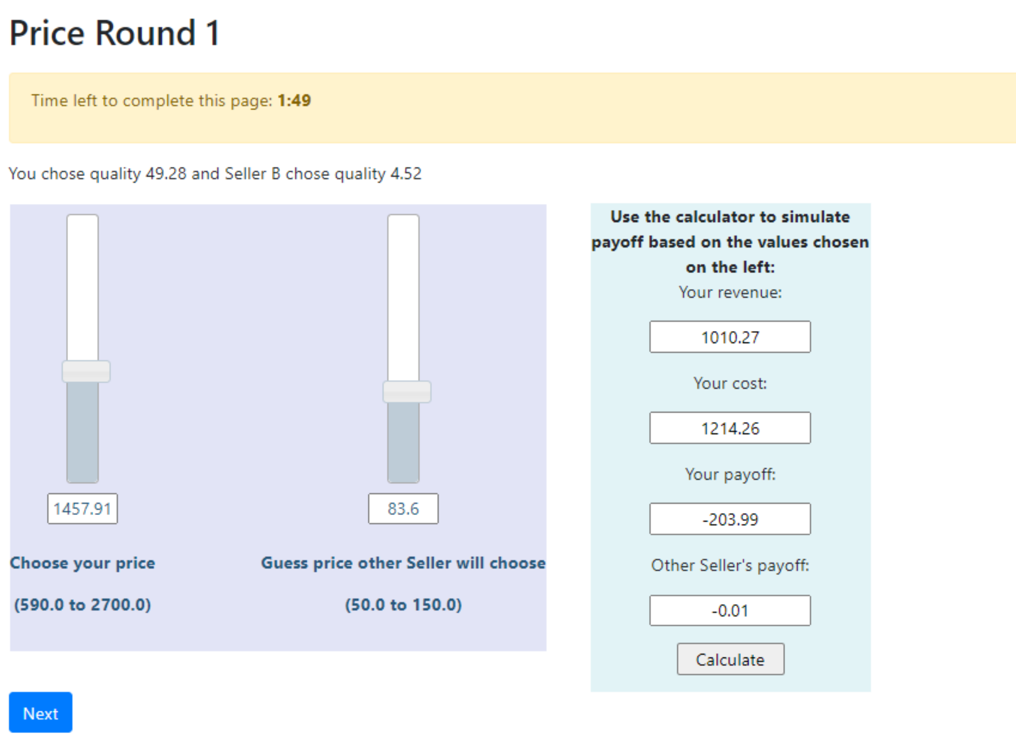
- **Seller B quality choice:** Seller B will next observe Seller A’s quality choice and will choose the quality of its product. Additionally, we ask Seller B to a) make a guess for the price Seller A will choose next ; b) choose the price they will choose based on the guess made in (a). The screen in which Seller B makes the choices is depicted below. To make choices, click on one of the options.



The calculator indicates the payoff using Seller B’s chosen values and the quality Seller

A chose earlier. To use the calculator, make the choices, then click the button “Calculate”. Click the button “Next” when ready to commit to the choices. If you do not press “Next” before the time to complete this page, the computer will randomly make a selection for you. But there is no need to rush your decisions, as there is plenty of time to choose. While Seller B makes the choices, Seller A waits. Seller A will not be able to see Seller B’s choices while waiting.

- **Price Choice:** Seller A and Seller B will observe the quality choices made by them and must set a price for their product. Additionally, we ask each player to make a guess for the other seller’s price, as shown below. To make choices, click on one of the options..



The calculator indicates the payoff with the price values that the seller is choosing now and the quality values previously chosen by Seller A and Seller B. To use the calculator, make the choices, then click the button "Calculate". Click the button "Next" when you want to commit to your price choice. If you do not press "Next" before the time to complete this page, the computer will randomly make a selection for you. But there is no need to rush your decisions, as there is plenty of time to choose.

Examples

Table 6 shows some examples of payoff under different combinations of qualities and prices.

Here is how you read the table. In line 1, the table shows that Seller A chose quality 24.5 in its quality round; this was followed by Seller B choosing quality 3.5 on its quality round; then, Seller A chose 913 as price for its product and Seller B chose 128 as the price for its product. As a result of all these choices, Seller A's revenue was 571.71 and its cost was 300.1. This resulted in a Payoff of 271.59 for Seller A. The combination of quality and price choices led Seller B's revenue to be 1.04, and its cost to be 6.1. This resulted in a Payoff of -5.09 for Seller B. You can check other combinations in Table 6.

Notice that there are cases in which the combination of qualities and prices in a round leaves you with negative payoff.

Notice that a Seller's cost depends on the level of quality that the Seller offers, and it is independent of the number of buyers that buy the product. Table 2 shows such relationship for Seller A and Seller B. For example, if Seller A chooses quality 19, Seller A's cost in a round will be 180.50. Similarly with Seller B: if it chooses quality 4 (see the second line, third column of Table 2), its cost will be 8.

Table 6: Payoff for Seller A and Seller B under different quality and price combinations

Choices				Results					
Quality Seller A	Quality Seller B	Price Seller A	Price Seller B	Revenue Seller A	Cost Seller A	**Payoff Seller A**	Revenue Seller B	Cost Seller B	**Payoff Seller B**
24.5	3.5	913.0	128.0	571.71	300.1	**271.59**	1.04	6.1	**-5.09**
24.5	3.5	913.0	129.0	572.15	300.1	**272.02**	0.61	6.1	**-5.51**
23.5	3.5	913.0	129.0	555.10	276.1	**278.98**	3.02	6.1	**-3.10**
23.5	4.5	913.0	129.0	536.27	276.1	**260.14**	16.25	10.1	**6.12**
23.5	4.5	2589.0	129.0	0.00	276.1	**-276.13**	129.00	10.1	**118.88**
23.5	4.5	2589.0	54.0	0.00	276.1	**-276.13**	54.00	10.1	**43.88**
23.5	5.5	2589.0	54.0	0.00	276.1	**-276.13**	54.00	15.1	**38.88**
37.0	5.5	1037.0	129.0	738.08	684.5	**53.58**	6.93	15.1	**-8.20**

Table 7: Costs under different choices

Quality Seller A	Cost Seller A	Quality Seller B	Cost Seller B
19	180.50	3.5	6.13
23.5	276.13	4	8.00
24	288.00	4.5	10.13
24.5	300.13	5	12.50
37	684.50	5.5	15.13

How are you going to be paid?

Out of the 10 effective trading rounds, the computer will select 4 effective paying rounds. Neither you nor the experimenter know which rounds are effective payment before the end of the experiment, as effective payment rounds are determined by the computer purely by chance. We will sum the earnings from the 4 effective payment rounds to determine your total number of points. Your total payment consists of the sum between your show up payment, the right questions you got from the quiz, and your effective payment from the experiment. If your earnings during the effective round are negative, your total payment will consists only of your show up fee and the result from the initial quiz. The experiment pay range varies between $5 and $35.

By the end of the experiment, the experimenter will pay you in private. You will be asked to wait outside and the experimenter will handle your money in an envelope. You may leave after you are paid.

Time to start the experiment

We will begin the experiment now. If you have any questions, raise your hand and the experimenter will go to you.

C Market outcomes: regression results

Tables 8 and 9 present regression results that further confirm the difference in means discussed above. Different regression models use the observed level of high quality, low quality, high-quality price, and low-quality price as the dependent variable. Again, these models compare the different treatments with the observed qualities and prices obtained in the benchmark (no misperception of qualities). The independent variables include a dummy taking the value of one if the observation belongs to treatment and zero otherwise, a conditional mean (intercept), and a time trend indicating the round during which subjects were making the choice. Subjects' demographic characteristics were added to balance the samples, but suppressed from the table as they offer no particular insight. The sign of the treatment dummy indicates the direction of the misperception treatment over the dependent variable, and the magnitude indicates the impact of quality supply or price charged under each treatment.

I start discussing the treatment effects of overvaluation of the high-quality product. The sign of the treatment effect on quality and price of the high-quality seller is as expected by theory, but the magnitude is lower. This is most likely due to the higher-than-expected quality offered under the benchmark case. Sign and magnitude of the treatment effect on high-quality prices have the expected sign and magnitude. Overvaluation of the higher quality has no significant impact on either quality offered or price charged by the low-quality seller. Similarly, undervaluation of the high-quality product produces the expected sign and magnitude of the treatment effect on quality and price of the high-quality seller, but no significant effect on the low-quality seller's choices.

Misperception of the low-quality product produces multiple forces. When consumers overvalue the low-quality product, the high-quality seller lower quality supplied (as expected), but at a higher magnitude than predicted by the theory. While no effect was detected

Table 8: Treatment effects under misperception of the high-quality product

	Overvaluation of high-quality			
	Quality, high	Price, high	Quality, low	Price, low
Const	29.59***	1393.6***	6.180***	123.0***
	(10.49)	(9.68)	(4.09)	(4.30)
Treat	3.003***	603.4***	-0.312	-10.99
	(4.00)	(25.18)	(-0.88)	(-1.75)
Round	0.106	17.82*	-0.0479	0.235
	(1.23)	(2.07)	(-1.00)	(0.41)
N	240	240	240	240

	Undervaluation of high-quality			
Const	35.45***	1637.4***	8.281***	139.0***
	(10.49)	(10.33)	(6.37)	(5.64)
Treat	-5.830***	-372.0***	0.402	-0.325
	(-3.88)	(-6.34)	(0.67)	(-0.06)
Round	-0.00380	0.0278	-0.0647	0.0420
	(-0.03)	(0.00)	(-0.54)	(0.08)
N	230	230	230	230

*** prob. < 0.01, ** prob. < 0.05. Models estimated using hierarchical random effects (at the session level). Standard errors clustered at the session level. Subjects' demographic characteristics included in all specifications.

Table 9: Treatment effects under misperception of the low-quality product

	Overvaluation of low-quality			
	Quality, high	Price, high	Quality, low	Price, low
Const	33.57***	1408.0***	5.336***	108.4***
	(12.55)	(7.73)	(3.52)	(4.64)
Treat	-3.124**	-137.2***	0.690	7.311**
	(-3.13)	(-4.69)	(0.69)	(2.10)
Round	0.0422	4.756	-0.0145	0.0647
	(0.30)	(0.82)	(-0.33)	(0.46)
N	250	250	250	250

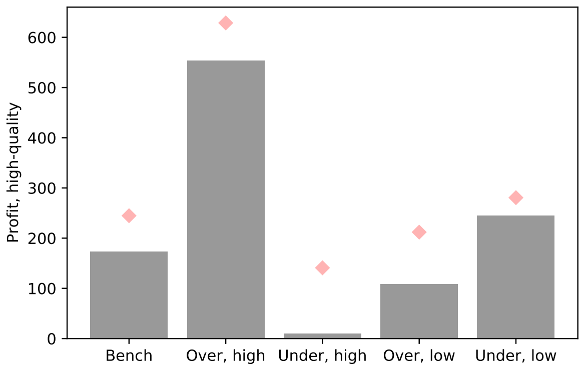
	Undervaluation of low-quality			
Const	33.21***	1576.3***	5.386***	105.9***
	(17.76)	(9.21)	(5.28)	(4.32)
Treat	0.0517	152.3***	-1.522***	-28.82***
	(0.09)	(5.70)	(-5.81)	(-6.72)
Round	-0.0745	-10.02	0.0196	0.0360
	(-0.88)	(-1.39)	(0.55)	(0.15)
N	230	230	230	230

*** prob. < 0.01, ** prob. < 0.05. Models estimated using hierarchical random effects (at the session level). Standard errors clustered at the session level. Subjects' demographic characteristics included in all specifications.

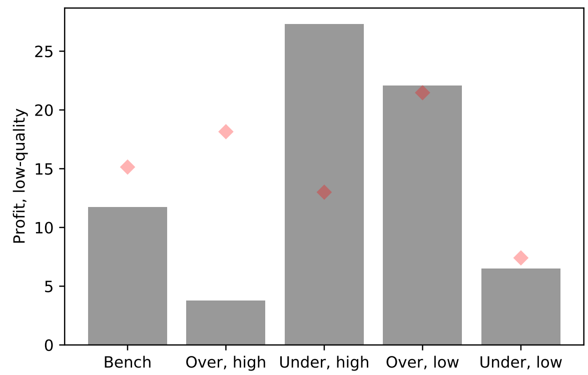
on supply of the lower quality (which is already higher than theory would predict, as discussed), overvaluation of the low-quality product allowed low-quality sellers to increase prices charged, with expected magnitude, as shown by the significantly positive treatment effect coefficient. Finally, the treatment effect of undervaluation of low-quality has no effect on the higher quality, but it increases the high-quality price in the expected magnitude; it also decreases the lower quality at the expected magnitude and sign. Finally, the treatment effect on low-quality prices is not as negative as expected by theory.

D Welfare outcomes: central tendency results

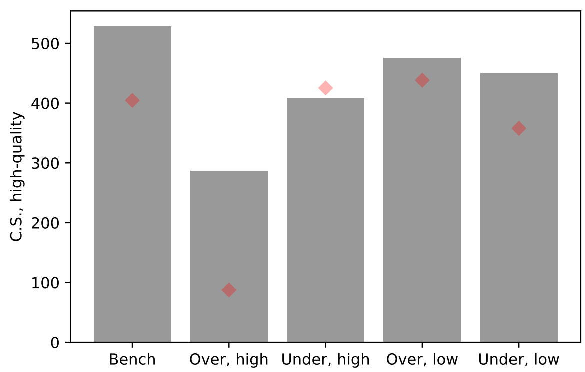
I show welfare outcomes calculated via a central tendency of the choices made during the experimental sessions. Specifically, I plug the averages of the qualities and prices of the high- and low-quality products on equations 1, 2, 3, and 4 to evaluate surplus measures before discussing treatment effects. Figure 5 shows the results.



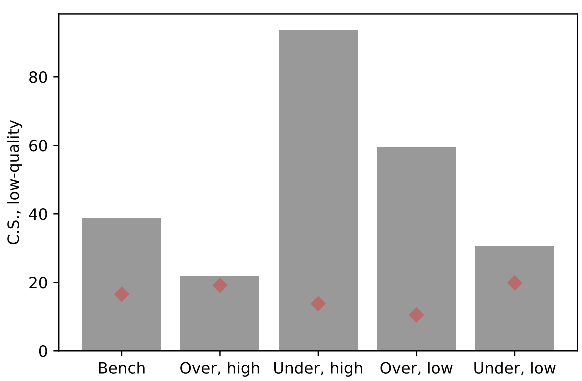
(a) Profit of high-quality seller



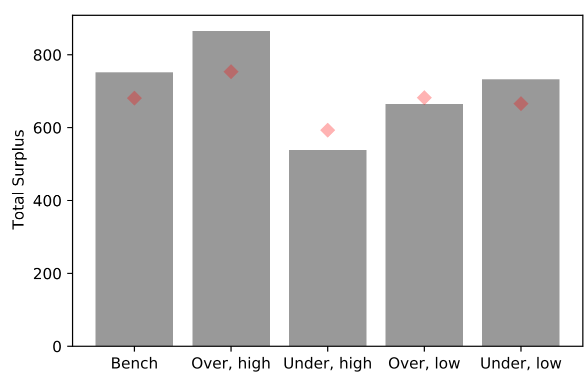
(b) Profit of of low-quality seller



(c) Consumer Surplus of high-quality product segment



(d) Consumer Surplus of low-quality product segment



(e) Total Welfare

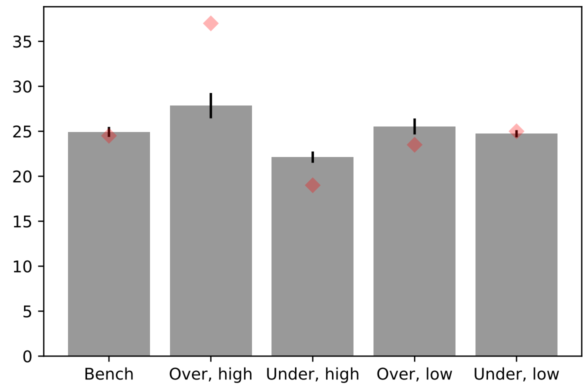
Figure 5: Welfare outcomes under different treatments. The height of the columns represent observed outcomes under average qualities and prices for each treatment, and the red marks represent theoretical equilibria

Holding prices and qualities on their observed averages produces total welfare outcomes close to the theoretical predictions, as shown by Figure 5e. But the distribution of surplus follows predictions only under some treatments. First, notice that high-quality profits and surplus of the high-quality segment are close to theoretical predictions (Figures 5a and 5c). Under overvaluation of high-quality, the profit of the low-quality seller is way below the predicted value; for undervaluation of high-quality, the profit of the low-quality seller is substantially above the prediction (Figures 5b). Surplus of the low-quality consumer is way above prediction, as shown in Figure 5d.

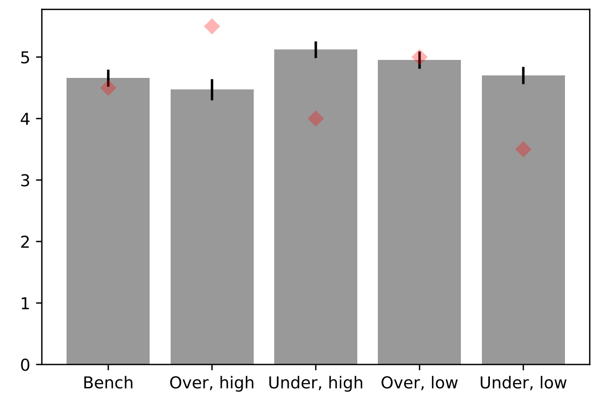
E Robustness check: discrete choice space

The session "Market Outcomes" in the main text alludes to the continuous space of quality and price choices as a possible source of sub-optimal behavior for low-quality sellers. To test whether the continuous space of quality and prices drove low-quality choices, I conducted a version of the experiment in which quality and prices are discrete. The experimental procedures are the same as described in the main text but the set of choices is as follows: quality of the high quality product must be chosen between {19, 23.5, 24.5, 25, 37}; for the low-quality product choices are between {3.5, 4, 4.5, 5, 5.5}; the high quality price must be chosen between {598, 913, 1037, 1155, 2589}; and for the low quality prices choices are {54, 80, 101, 128, 129}.

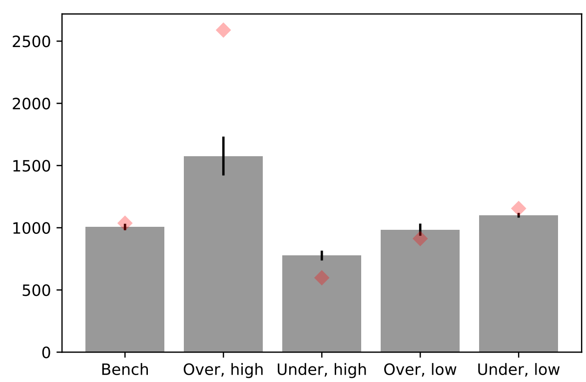
Market outcomes largely follow the same pattern seen in the main text, as seen in figure 6. Average high quality are close to the optimal value and is highly responsive to misperception. Choices for low quality are closer to predicted behavior, but prices are still far from what the model predicts, again suggesting a strategical failure from subjects playing the low-quality seller. Quality-adjusted prices are still too low compared to what the theory predicts. The results were obtained



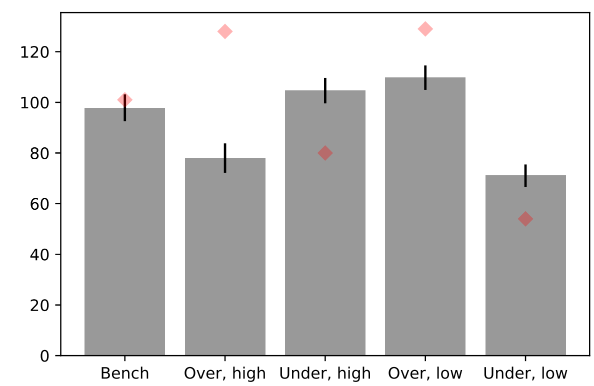
(a) Quality of high-quality product



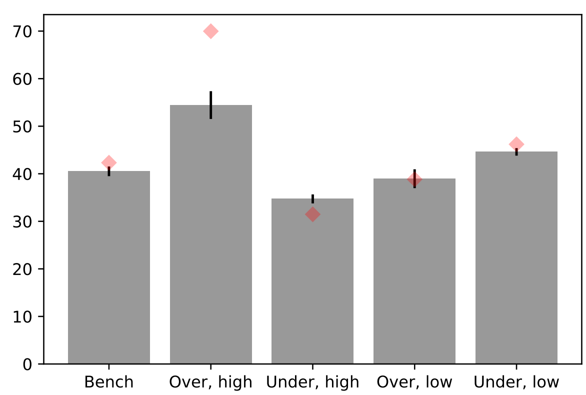
(b) Quality of low-quality product



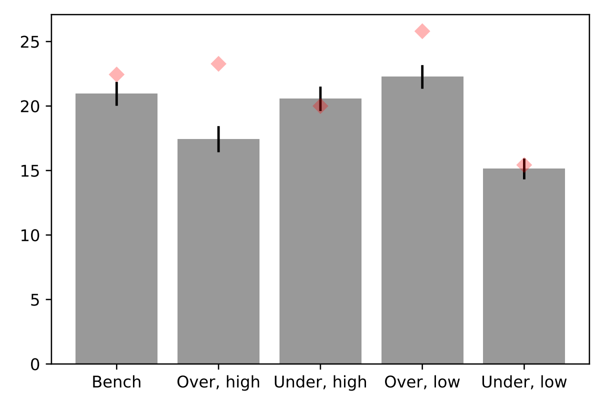
(c) Price of high-quality product



(d) Price of low-quality product



(e) Quality-adjusted prices of the high-quality product



(f) Quality-adjusted prices of the low-quality product

Figure 6: Market outcomes under different treatments (discrete case). The height of the columns represent average observed outcomes, the red marks represent theoretical equilibria, and the bars are the 95% C.I.