CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Unfair commercial practices in a pit market: evidence from an artefactual field experiment

Francesco Bogliacino[1]* (iD), Rafael Charris[2], Cristiano Codagnone[3], Frans Folkvord[4], Felipe Montealegre[5] and Francisco Lupiáñez-Villanueva[6]

[1]Università degli Studi di Bergamo, Bergamo, Lombardia, Italy, [2]Economic Science Institute, Chapman University, Orange, CA, USA, [3]Dipartimento di Scienze Sociali e Politiche, Università degli Studi di Milano, Milano, Lombardia, Italy, [4]Tilburg School of Humanities and Digital Sciences, Tilburg University, Tilburg, Noord-Brabant, Netherlands, [5]Dipartimento di Scienze Economiche, Università di Bologna, Bologna, Emilia Romagna, Italy and [6]Faculty of Information and Communication Science, Universitat Oberta de Catalunya, Barcelona, Catalunya, Spain

***Corresponding author**: Francesco Bogliacino, email: francesco.bogliacino@unibg.it

Abstract

Commercial practices such as drip pricing, reference pricing and best-price guarantee can be used to set higher prices and mislead consumers, but protective measures can restore efficiency. In a placebo-controlled market experiment, we examined a treatment allowing for the use and misuse of commercial practices. Three additional treatments tested the effects of formal sanctions, informal sanctions and a regret nudge. We found that commercial practices led to higher prices, cheating was systematic and regret nudging was ineffective. Furthermore, formal and informal sanctions reduced both the likelihood of using commercial practices and the likelihood of cheating, leading to welfare increases.

Introduction

When companies bring their products to the market, they adopt various commercial practices in an attempt to promote sales, from making prices more salient to changing the framing of the offers. These practices are a vehicle for the provision of information but, it is argued, can also limit competition, confuse consumers and clutter the allocative outcome of the market. Thus, in this study, we pose three research questions to shed light on the consequences of commercial practices: Do commercial practices lead to higher prices? Do sellers cheat by providing misleading information when using such practices? Can protective measures re-establish efficiency?

We focus on three practices, chosen because they are widespread in online services platforms: drip pricing (DP), reference pricing (RP), and best-price guarantee (BPG). DP is the practice of decomposing a price into multiple components that are presented sequentially to buyers. RP presents the current price as a discounted price (usually by

displaying the price alongside a higher price that has been struck through). A BPG is a statement alleging that the current price is the best price available in the market.[1]

We also focus on three protective measures. The first involves the random auditing of sellers and the imposition of fines if they are found to be cheating (formal sanctions). The second is a decentralized audit in which buyers can collectively and informally sanction a seller for cheating. The third is a nudge that leverages individual regret to induce more precautionary behaviour by buyers. These three measures are not specific remedies in response to particular practices and are readily scalable.

We used an experimental approach to circumvent the difficulty of identifying plausible exogenous variations in both commercial practices and protective measures in natural settings.

In our experiment, participants made decisions in the role of either a seller or a buyer over five trading periods regarding tourist packages to various destinations. Sellers posted their prices, and buyers simultaneously approached the sellers under time pressure. In one treatment, sellers were able to endogenously choose from a menu of commercial practices (DP, RP and BPG), which could be used either honestly or deceptively. Three additional treatments were used to test our remedial actions. In the treatment with formal sanctions, sellers were probabilistically audited and fined if they were found to have cheated. In the treatment with informal sanctions, buyers were able to respond to a deceptive practice by sending a signal. If at least three buyers sent a signal in relation to a particular seller, the seller was fined. Finally, in the regret nudge treatment, buyers received a feedback emoticon before confirming a transaction, a happy face if they were paying less than the average price and a sad face otherwise.

The results showed that when using DP, RP and BPG, sellers post higher prices and cheat, thereby misleading consumers and reducing efficiency. The evidence regarding the effectiveness of the protective measures was mixed. Formal sanctions were effective, regret nudging was not effective and informal sanctions were only effective to a limited extent. The protective measures also produced other qualitative effects, with sellers responding by shifting from more easily detectable practices (such as a best-price guarantee) to less easily detectable alternatives (such as DP).

The remainder of the paper is organized as follows. Section 'Policy background and related literature' presents the policy background and a literature review, section 'Experimental design' presents the experimental design, and section 'Results' presents the results. Section 'Discussion and conclusions' discusses the results and presents concluding remarks. Additional information is presented in Supplementary Appendix.

Policy background and related literature

The unfair commercial practice directives and a behavioural perspective

Companies use a variety of techniques in an effort to increase sales. In legal terms, these techniques are referred to as commercial practices. In Europe, they are regulated

---

[1]In the real world, a best-price guarantee is a best-price claim but also commits the seller to reimbursing the difference to the buyer if a cheaper price is found elsewhere in the market. Because of the nature of our experiment, this second element of the practice could not be implemented, and thus, we focused on the limited effect of the claim to offer the lowest price. Since we used the label *best-price guarantee* in the experiment, we maintain it here.

by the Unfair Commercial Practice Directive, whose rules are broad and technologically neutral, applying equally to online platforms and brick-and-mortar stores. The regulation of commercial practices based on this perspective adopts the perspective of the average consumer, who is defined as sufficiently circumspect and paying enough attention (European Parliament and Council, 2019). This regulation aims to prevent sellers from crossing the line from persuasion to manipulation, and a necessary condition is material effectiveness. In Europe, material effectiveness is interpreted via an empirical transaction test: the consumer agrees to a deal that would not have occurred otherwise.[2]

Both pillars – technological neutrality and the average consumer – have come under increasing scrutiny and scepticism with the growing use of cognitive psychology to explain consumer behaviour (Thaler, 1980) and the increasing presence of e-commerce and digital platforms (Calo, 2014; Jabłonowska *et al.*, 2018; Mathur *et al.*, 2019).

First, if consumers were consistent, choice architectures based on complex information or easily verifiable lies would not condition their decisions. To the contrary, bounded rationality can generate an entire set of market failures (Hanson & Kysar, 1999). Indeed, if public agencies can exploit cognitive vulnerabilities to nudge decision-making (Thaler & Sunstein, 2003, 2008), platforms can learn to do the same, but for profit.

Second, when consumers buy offline, the law gives ample contractual freedom to the parties, provided that there is agreement on the terms of the contract (consent). The consent is meaningful in offline environment because (1) the request for consent is infrequent, (2) the risk of giving consent is clear and (3) subjects have an incentive to take each request seriously. On the contrary, these three conditions are consistently violated in the online environment (Richards & Hartzog, 2019).

The popularity of behavioural economics and the problems associated with online privacy and consent may have revealed the limits of the existing regulation of commercial practices, but the truth is that the significant body of case law that has developed around deceptive communication and unfair commercial practices signals an increase in problematic behaviour by companies and an associated increase in litigation (Keller, 1996; Stuyck, 2015; Mathur *et al.*, 2019; Netherlands Authority for Consumers & Markets, 2020; Luguri & Strahilevitz, 2021).

Most of the legal debate internalizes a demand approach, centred on eliciting consumers' response to specific practices (Huck & Wallace, 2015), thereby limiting the scope of the analysis. Industrial economists have paid more attention to the supply side, studying the consequences of the use of shrouding and complexity as anti-competitive tools (Kalayci & Potters, 2011; Gu & Wenzel, 2015, 2020; Kalaycı, 2016; Crosetto & Gaudeul, 2017; Normann & Wenzel, 2019; Rasch *et al.*, 2020). However, this body of literature continues to overlook the endogenous choice of practices from a menu of options (instead of a simple extensive margin) and misleading use (instead of simple adoption). A direct consequence of this limited attention to the

---

[2]In the US, the unfair commercial practices are regulated by the Federal Trade Commission Act. The transaction test in the US is interpreted as the inducement of a net loss, that is, the benefits do not adequately compensate the buyer for the price paid (Luguri & Stahilevitz, 2021).

endogenous choice of options is the paucity of proposed protective measures, which are usually limited to either prohibition (Rasch *et al.*, 2020) or increasing market entry (Kalaycı, 2016).

Nevertheless, providing misleading information represents a form of cheating, and thus public and experimental economists have studied this behaviour and ways to discourage it (van de Ven & Villeval, 2015; Celse *et al.*, 2019; Gneezy & Kajackaite, 2020; Irlenbusch *et al.*, 2020). Adapting an argument by Becker (1968), the likelihood of detection and severe punishment discourage cheating. A centralized audit with sanctions constitutes a standard application of this principle.

Formal sanctions are not the only mechanism available. Studies of social dilemmas have shown that in the presence of incentives to free ride, a demand for group norms leads to the establishment of informal sanctions including peer pressure, gossip and ostracism (Homans, 1961; Coleman, 1990). Informal sanctions are usually effective in laboratory and laboratory-in-the-field experiments (Masclet *et al.*, 2003; Lopez *et al.*, 2012).

Some studies have found that emotions can either usher in or undermine dishonest behaviour. For example, shame and shame avoidance are important in explaining conformity and social cohesion (Lewis, 1992; Coricelli *et al.*, 2014), which is why cheating can be mitigated by scrutiny. Similar considerations apply in relation to guilt and guilt aversion (Charness & Dufwenberg, 2006; Battigalli & Dufwenberg, 2007; Ellingsen *et al.*, 2010). In an effort to enhance the existing body of literature, we focus on regret (Loomes & Sugden, 1982; Coricelli *et al.*, 2005). We take a different angle, though. In our design, we do not manipulate incidental emotions, but cognition about emotions (Rick & Loewenstein, 2008). In particular, we manipulate the possibility that participant feels regret at the moment of confirming the purchase and we make salient that this may occur, to indirectly shape the belief of forward-looking sellers. The use of an emoticon that switches expression below and above the average behaviour has featured in previous policy interventions (Allcott & Mullainathan, 2010) and we decided to test it in our setting. With respect to the original application in energy consumption, the emoticon in our setting is also a way to pause the purchase behaviour forcing the buyer to revaluate the decision (Dworking, 1988). In fact, it is embedded into a confirmation page.

The background of this study

The experiment designed for this article constitutes part of a larger study providing evidence-based recommendations for consumer protection in the e-travel sector. The e-travel sector comprises online travel booking – hotels, vacation rentals and packaged holiday – and mobility services (flights, car rentals, transportation services such as trains and buses and ride-sharing services made available through platforms and applications).

E-travel has been leading other industries in the penetration of online platforms and digital transformation and affecting the way to do business and the use of commercial practices. The digital landscape combines information overload on the side of consumers and increasing tools in the hands of companies. The new tools raised an alarm and new regulatory proposals, since ensuring a high level of consumer trust is

perceived as fundamental for the functioning of the market. However, most of the discussion revolves around a proper test of deception. The regulation both in the EU and the US requires a combined assessment that the practice misleads and changes behaviour (Sahni & Nair, 2020; Luguri & Strahilevitz, 2021).

This study offers an alternative angle. Instead of trying to envisage a test that disentangles fair from unfair use of practices within transaction data, we explore a direct approach. In a controlled environment, we can observe the provision of correct and incorrect information to consumer, identifying deception. Additionally, in an environment with supply and demand, we can directly quantify welfare comparing it with the benchmark of a competitive market. This overcomes the need to jointly test deception and effectiveness, as we can compare outcomes with and without commercial practices, under a normative criterion of efficiency.

Finally, while a demand approach only allows testing the prohibition of practices and counteracting measures which are practice specific, for example, forms of visual cues or nudges such as our emoticon treatment, our market environment allows to test regulatory interventions that are either centralized or decentralized (as our two additional treatments of formal and informal sanctions), which are easily scalable.

Experimental design

Design overview and treatments

We conducted an artefactual field experiment (Harrison & List, 2004), in which the participants were either sellers or buyers in a pit market where sellers were given one chance to update their prices and buyers approached the sellers simultaneously rather than sequentially (Smith, 1962; Davis & Holt, 1993). Roles were randomly assigned at the beginning of the session and maintained throughout the experiment. Market sessions consisted of five trading periods, in which the participants negotiated tourist packages to five European capitals (Rome, Vienna, Paris, Madrid and Berlin).

Each participant was given 25 experimental currency units (ECUs)[3] at the beginning of the session, and there were five varieties of a framed good on offer in the market. Sellers had the right to sell one unit of a variety of a good and were told its cost, which represented the minimum price that they were willing to accept. Buyers could purchase one unit of any of the varieties of the good and had private valuations of all varieties, that is, the maximum price they were willing to pay for each variety. The cost and valuation parameters were private information. In contrast to the standard pit market, the sellers were price-makers and orders were closed on a first-come-first-served basis.

Transactions occurred as follows. Sellers posted a price, after which buyers could post an order for a specific package. The transaction was closed after the first buyer posted a purchase order. The seller earned the difference between the price at which the transaction was closed and the cost, while the buyer earned the difference between their valuation and the price at which the transaction was closed. In cases where the transaction did not occur, earnings were zero.

---

[3]The initial endowment ensures the satisfaction of the participation constraint of subjects from the general population. Additionally, it covers the costs of acquiring information for sellers, or covers the losses from fines, etc.

This was a between-subject design with five experimental conditions randomized at the session level: (1) *Control*: baseline task, 10 sellers and buyers per session, five possible packages to trade; (2) *Commercial Practice*: similar to the Control treatment, but the sellers had access to commercial practices and could cheat (details are presented in the next subsection); (3) *Formal Sanctions*: similar to Commercial Practice, but with a 20% probability of sellers being audited and fined if found to be cheating (with the buyers sharing any money obtained through fines); (4) *Informal Sanctions*: similar to Commercial Practice, but buyers could simultaneously send a signal to other buyers denouncing a seller. If at least two additional buyers sent a similar signal (for a total of three), the seller incurred a sanction, the proceeds of which were split among the plaintiffs; (5) *Regret*: after posting a successful order, the buyers should confirm the purchase. When asked to confirm, the buyer saw a smiling emoticon if the price of the transaction was below the average market price or a sad emoticon otherwise.

Data were collected from laboratory experiments in two countries, Spain and Germany, and the protocol was programmed in oTree (Chen *et al.*, 2016). Participants, who were all aged between 25 and 40, were recruited via a market research company. The company randomized invitation from a panel. No additional quotas were included beside age. The experiment took place in research facilities provided by the research company.

Payments were in ECUs that could be exchanged for real currency at a predetermined exchange rate of 1 ECU = 1 EUR in Spain and 1 ECU = 1.2 EUR in Germany to control for differences in purchasing power parity. At the end of the session, we randomly selected one round on which payments were based to eliminate any income effects. The initial endowment represented a participation fee.

Procedure

The research assistant leading the session read out the general instructions, while the participants were asked to follow the instructions on the screen. At the beginning of round one, the participants were randomly assigned to be either a buyer or a seller and maintained the same role across all five rounds of the experiment.

The value (V) and cost (C) sets from which we generated demand and supply were $V = \{100, 100, 90, 80, 80, 70, 60, 60, 50, 40\}$ and $C = \{10, 10, 20, 30, 40, 40, 50, 50, 60, 70\}$ for each destination. The specific demand and supply were generated by assigning five valuations, one per package, to every buyer, and one cost for one specific package to every seller. The procedure was a random draw without replacement.

Sellers moved first by posting an initial price and then deciding whether to bear a cost (1 ECU) to acquire information on the initial prices and destinations offered by competitors. Then, they could either confirm or revise their initial proposal and post a final price. When revising their prices, in all treatments other than the Control treatment, sellers could choose one of three commercial practices: BPG, RP or DP. BPG claimed that the package on offer was the lowest priced among all packages to the same destination. RP suggested that the seller's final price offered a discount on the initial price. DP enabled the seller to display a price for the package net of

the cost of acquiring information. The difference between the real price and the displayed price was the cost of viewing the full list of packages, regardless of whether the seller decided to pay this cost or not.

The seller was able to cheat when adopting these practices. Cheating took the following forms: (a) the seller offered a best-price guarantee, but there was a lower price on offer in the market; (b) the seller misreported the discount with respect to the original price; and (c) the seller included the cost of acquiring information without having been charged for it. The existence of the three practices and the possibility of cheating was common information.

When it was their turn, buyers saw a list of all the offers posted by the sellers and could place an order to buy one package. If two buyers placed orders for the same package, the transaction was closed on a first-come-first-served basis.

The process differed in relation to the other treatments in the following ways. In the treatment involving formal sanctions, sellers were randomly audited and fined if they were found to have cheated, with buyers sharing the proceeds of the fine. In the treatment involving informal sanctions, a buyer could report a seller for cheating at a cost of 1 ECU, and if at least two other buyers also reported the same seller for cheating, the seller was fined, regardless of whether the seller had actually cheated.[4] The fine was calculated as 2 ECUs multiplied by the number of plaintiffs and was divided equally among the plaintiffs. In the treatment with regret, the buyer was given a chance to confirm the purchase and saw a message with an emoticon prior to do that, a smiling emoticon if the price of the transaction was below the average market price or a sad emoticon otherwise.

Participants received the following feedback after each round: whether they had sealed a transaction and the related pay-off (buyers and sellers in all treatments), whether they had been audited and fined (sellers in formal sanctions), whether they had been reported and fined (sellers in informal sanctions) and whether they had paid more or less than the average price paid (buyers in the regret treatment).

Each session lasted for about 90 minutes, with minimal variation.

The sample was balanced by gender (51.25% were male), the average age was 31.79 yr with a standard deviation (SD) of 6.63 yr, 31.75% of participants held a university degree and 30.75% reported having a family income between 20,001 Euros and 40,000 Euros. There were 20 participants in each session (10 buyers and 10 sellers), and two sessions were run in each country (Spain and Germany) for each of the five treatments. Given the five round and that we are pooling together the two countries, this is equivalent to 200 demand observations and 200 supply observations, per treatment (1000 in total).[5] Data were collected in December 2019.

---

[4]Note that this creates a coordination game with multiple equilibria. Experimental evidence suggests that focal points guide the participants' behavior in this class of game (Mehta *et al.*, 1994). We conjecture that this should make coordination easier in relation to genuine cheaters.

[5]The number of subjects was constrained by the budget, but a power analysis conducted for the price variable, which was the main outcome, showed that the study was properly powered. In fact, assuming the equilibrium price of 60 as mean outcome in the control, a two-sample *t*-test with a 5% alpha and a (small-to-medium) minimum detectable effect of 10 units (one third of a SD, assuming a common SD of 30) returns a power of 90%.

Hypotheses

We tested four hypotheses based on the review of the literature in the previous section.

H1: Commercial practices lead to higher prices.
H2: Sellers systematically cheat when using commercial practices.
H3: The introduction of protective measures reduces cheating. Additionally, sellers endogenously adopt cheating practices where they are less likely to be detected and sanctioned.
H4: The introduction of commercial practices reduces welfare but responding with protective measures increases welfare.

Hypothesis 1 is based on the observation that commercial practices are used to differentiate a seller's goods from those of their competitors, thereby introducing market power. Hypotheses 2 and 3 are inferred from the assumption that people respond to incentives. A fine in the treatment with formal sanctions is more likely to occur and is more costly, supporting the view that formal sanctions are more effective than informal sanctions and regret. In stating our H3, we are assuming that BPG is more salient than the other two practices. Under time pressure to place an order and seal a transaction, practices that imply comparison are immediately scrutinized. This is consistent with the theory of salience (Bordalo *et al.*, 2013). RP introduces a different type of comparison, ex ante vs ex post (because of the strike-through price), instead of that across options. DP stands in the middle: by showing only a partial price, can promote price comparison but by requiring the calculation the full price may be more difficult to process under time pressure.

Finally, Hypothesis 4 addresses the welfare implications of Hypotheses 1–3.

Results

Analysis of pricing behaviour

We present the results of our testing of Hypotheses 1–4 in sequence. Column (1) in Table 1 shows the results of ordinary least squares (OLS) regression of posted prices on the dummies for treatments using round, gender, age, education and country as controls, and with standard errors clustered at the session level, the level at which we assigned the treatments. Controls are elicited in the post experimental questionnaire and include only variables unlikely to be an outcome affected by treatment and whose inclusion increases precision of the estimates. Column (2) controls for the cost parameters assigned to sellers. We also present a nonparametric analysis of this outcome variable in Supplementary Appendix. This also applies to all other outcome variables analysed in this section.

The mean price in the control, as reported in the Table 1, is 58.13 ECUs. Using the results from Column (2), compared with the control condition, sellers posted prices that were 11% higher (6.71 ECUs, $t = 2.37$, $p = 0.028$) when allowed to use commercial practices. The threat of formal sanctions did not reduce prices, because the prices posted under this treatment were 14% (8.27 ECUs) higher than those under the

Table 1. Outcome variables: price posted and willingness to buy

Variables	(1) Price	(2) Price	(3) WTB	(4) WTB	(5) WTB
Commercial practices	6.25*	6.71**	−0.04	−0.04	−0.04
	(3.32)	(2.83)	(0.03)	(0.03)	(0.03)
Formal sanctions	7.45*	8.27	−0.04	−0.04	−0.04
	(4.06)	(4.90)	(0.03)	(0.03)	(0.03)
Informal sanctions	3.90	4.83	−0.20***	−0.20***	−0.21***
	(2.84)	(3.00)	(0.04)	(0.04)	(0.04)
Regret	3.39	5.39**	−0.02	−0.03	−0.03
	(2.29)	(2.45)	(0.03)	(0.03)	(0.03)
Constant	67.12***	35.79***	0.80***	0.86***	0.91***
	(7.26)	(6.15)	(0.07)	(0.10)	(0.14)
Observations	1,000	1,000	1,000	1,000	1,000
R^2	0.03	0.26	0.22	0.22	0.22
Controls	Yes	Yes	Yes	Yes	Yes
Mean outcome (control)	58.13	58.13	0.960	0.960	0.960
Cost parameters		Yes		Yes	Yes
Valuations					Yes

Notes: Clustered standard errors (at the session level) are shown in parentheses. *p < 0.1, **p < 0.05, and ***p < 0.01. OLS regressions. Controls include round, gender, age, education, country, and individual seller valuation.

baseline treatment ($t = 1.69$, p = 0.108). The 8% (4.83 ECUs) difference in price between informal sanctions and control was not statistically different ($t = 1.61$, p = 0.124). The regret treatment had prices that were 9% (5.38 ECUs) higher and statistically different from the control ($t = 2.20$, p = 0.041). As a result, the null hypothesis in H1 was rejected in favour of the alternative hypothesis.

Before posting prices, sellers could spend money to acquire information about their competitors. They did it on average 42% of the time. Differences across treatments were not statistically significant. In the absence of commercial practices, the likelihood of acquiring information was 39%. When commercial practices were introduced, this probability was 42% ($t = 0.39$, p = 0.703). In the presence of protective measures, it increased to 43.5% in the treatment with formal sanctions ($t = 0.51$, p = 0.618) and 47.5% in the treatment with informal sanctions ($t = 0.80$, p = 0.432), while it decreased to 38% in the treatment with a regret nudge ($t = −0.59$, p = 0.563).

Analysis of willingness to buy

Column (3) in Table 1 shows the estimated coefficients from an OLS regression of the willingness to buy over the treatment dummies and the usual controls, which were added for precision. To control for prices but overcome endogeneity, Column (4)

controls for the average cost parameter in each round (which was randomly assigned) and Column (5) for both average cost parameter and individual valuations. We use this last column as reference. On average, buyers placed an order 96% of the time in the control. In the presence of commercial practices, willingness to buy was not statistically different from the control ($t = -1.29$, p = 0.213). Similarly, the average willingness to buy was not different from the control in the treatment with formal sanctions ($t = -1.58$, p = 0.131) and in the treatment with a regret nudge ($t = -0.95$, p = 0.355). The willingness to buy faced a reduction by 20 percentage points (pp) in the treatment with informal sanctions ($t = -5.18$, p < 0.001).

This finding should not be interpreted as ineffectiveness of the practices or the measures. In fact, all consumers have at least one profitable deal in each round, as shown by the fact that close to 100% of the consumers place an order. Notice that this holds by construction: following standard protocols for pit markets, there exist mutual gains from trade. Conditional on prices, the likelihood to place an order does not change due to commercial practices because gains from trade are not wiped out. With an exception, of course, since behaviour did change in the treatment with informal sanctions.

Thus, we were left with the open question of why informal sanctions had such a pronounced effect on willingness to buy after we had controlled for price differences. One possibility is that instructing buyers to report cheaters might have created an additional cognitive cost, with buyers more sceptic and more focused in detecting cheating than finding profitable deals. As it may occur that the introduction of extrinsic motives crowds out intrinsic motives (Gneezy & Rustichini, 2000), similarly, in this setting, the introduction of the sanctioning role may have changed the consumers' perception of their 'behavioural script' (Bicchieri, 2006) from buyers to monitors.

Needless to say, to evaluate the allocative property of commercial practices and its comparison with respect to the benchmark of the competitive market, sealed transactions and not posted order matter. This welfare analysis is performed in the subsection 'Welfare effects of commercial practices and protective measures' below.

Analysis of cheating

In Table 2, we report the OLS regression of the use of practices and cheating on treatment dummies and control, clustering standard error at the session level. When allowed to do so, sellers adopted commercial practices 89.5% of the time. BPG accounted for 54% of the use, RP for 36% and DP for the remaining 10%.

The introduction of protective measures had a non-statistically significant effect on the use of commercial practices. Centralized auditing and informal sanctions reduced the use of commercial practices by 6 pp and 4 pp, respectively ($t = -1.61$, p = 0.128 and $t = -0.82$, p = 0.425, respectively), while regret resulted in a slight increase of 4 pp ($t = 1.63$, p = 0.123). The results of the regressions are presented in Column (1) of Table 2.

This negligible effect conceals a dramatic change in composition. To identify this change, we replaced the original outcome (where the dummy was equal to 1 if a practice was used) with three separate dummies, one per individual practice. By

Table 2. Outcome variables: likelihood of using a specific commercial practice and likelihood of cheating using a specific commercial practice

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Use	BPG (Use)	RP (Use)	RP (Use)	Cheat	BPG (Cheat)	RP (Cheat)	RP (Cheat)
Formal	−0.06	−0.23***	0.07**	0.10***	−0.16***	−0.23***	0.02	0.04
	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)	(0.05)	(0.02)	(0.03)
Informal	−0.04	−0.12**	−0.01	0.10***	−0.07	−0.12*	−0.02	0.06
	(0.04)	(0.06)	(0.04)	(0.03)	(0.06)	(0.06)	(0.03)	(0.04)
Regret	0.04	−0.13*	0.03	0.13	0.03	−0.11*	0.04	0.10*
	(0.02)	(0.06)	(0.10)	(0.08)	(0.05)	(0.06)	(0.06)	(0.05)
Observations	800	800	800	800	800	800	800	800
R^2	0.04	0.06	0.02	0.03	0.06	0.05	0.02	0.03
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome (control)	0.895	0.485	0.325	0.0850	0.625	0.435	0.135	0.0550

Notes: Standard errors are shown in parentheses. *$p < 0.1$, **$p < 0.05$ and ***$p < 0.00$. OLS regressions, with round, gender, age, education, country and individual seller valuation used as individual controls. Clustered standard errors are at the session level.
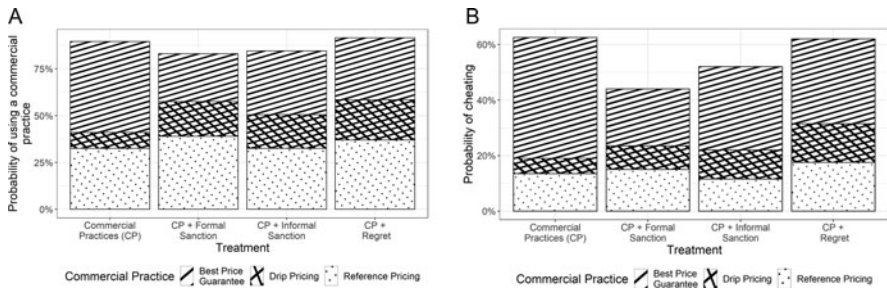
Figure 1. Probability of (a) using a commercial practice and (b) cheating by condition.

construction, summing the effect of a treatment on the use of the three practices is equal to the effect of the treatment on the initial outcome variable. The results of these additional regressions are presented in Columns (2)–(4) in Table 2.

Introducing formal sanctions decreased the use of BPG by 23 pp ($t = -5.17$, p < 0.001), whereas the use of RP increased by 7 pp ($t = 2.14$, p = 0.049), and the use of DP increased by 10 pp ($t = 3.28$, p = 0.005). Under informal sanctions, the use of BPG fell by 12 pp ($t = -2.21$, p = 0.043) and the use of DP increased by 10 pp ($t = 3.00$, p = 0.009), while the use of RP is unaffected ($t = -0.30$, p = 0.770) Finally, in the regret treatment, the use of BPG decreased by 13 pp ($t = -2.11$, p = 0.053), while the use of RP ($t = 0.31$, p = 0.764) and DP ($t = 1.66$, p = 0.118) is unaffected. Figure 1a plots these results showing how the use of commercial practices differed across treatments.

Sellers mostly used commercial practices to provide misleading information that worked to their advantage. This cheating occurred 62% of the time, largely supporting H2.

The likelihood of cheating decreased by 16 pp in the case of formal sanctions ($t = -3.72$, p = 0.002), 7 pp in the case of informal sanctions ($t = -1.24$, p = 0.233), and increase by 3 pp in the case of regret ($t = 0.63$, p = 0.539), with the latter two differences lacking statistical significance.

Columns (6)–(8) in Table 2 present the same information as Columns (2)–(4) but using cheating as the outcome. Formal sanctions made a seller 23 pp less likely to cheat using BGP ($t = -4.97$, p < 0.001), whereas the differences in cheating using RP ($t = 0.91$, p = 0.378) and DP ($t = 1.45$, p = 0.166) are not statistically significant. Informal sanctions reduced cheating using BGP by 12 pp ($t = -1.96$, p = 0.069) but the 2 pp difference in cheating on RP ($t = -0.53$, p = 0.606), and the 6 pp increase in cheating using DP ($t = 1.47$, p = 0.163) are not statistically significant. Regret reduced cheating using BGP by 11 pp ($t = -1.84$, p = 0.085), but increased cheating using DP by 10 pp ($t = 2.02$, p = 0.061). The 4 pp difference in cheating using RP is not statistically significant ($t = 0.79$, p = 0.441). Figure 1b shows how much cheating was accounted for by individual practices.

Thus, the alternative hypothesis for H3 is strongly supported. Profit provides the incentive for sellers to cheat, while protective measures force them to change the type of practices used. Possibly, this shift is not random but towards those practices that are safer, as explained above. The net effect was strongest for formal sanctions, as expected.

We make three final observations. First, users of DP posted the highest prices, followed by users of BPG. The average price in the control condition was 58.13 ECUs[6] (which was within the predicted equilibrium range), while non-users of practices and users of BPG, RP and DP sold at average prices of 61.14, 64.05, 61.29 and 64.66, respectively.

Second, on average, among closed transactions, sellers cheat 48.49% of the times, of these 55.90% used BPG, 26.09% RP and 18.01% DP. The proportions in the treatment without protective measures are 64.29%, 28.57% and 7.14%. These proportions become 50%, 34.38% and 15.63% in the treatment with formal sanctions, 65.79%, 21.05% and 13.16% in the treatment with informal sanctions, and 44.90%, 22.45% and 32.65%, in the treatment with regret.

Third, the rate of decentralized reporting in the informal sanctions treatment (10%) was about half the rate of centralized auditing (20%) in the formal sanctions treatment, with BPG triggering most reporting by buyers (19%). However, informal sanctions resulted in antisocial behaviour (Hermann *et al.*, 2008), with 40% of those reported found not to have cheated.

Welfare effects of commercial practices and protective measures

We conclude this section by exploring some of the equilibrium properties of the market. We first explain the procedure to calculate the level of efficiency and define the total surplus as the sum of the consumer surplus and the producer surplus (this corresponds to the area below the demand curve and above the supply curve). Because sellers and buyers can trade at most one unit and products are horizontally differentiated, compute the total surplus by assigning each available product to the highest evaluation. Then, divide the sum of gross profits from transactions for each party by the total surplus to obtain the level of efficiency.

Figure 2 shows the level of efficiency by treatment and by round. In the control condition, efficiency starts at around 40% and increases to 50%, most likely because of learning effects. This level of efficiency is comparatively low (Chamberlin, 1948; Smith, 1962; List, 2002; Baghestanian *et al.*, 2014). We speculate that the loss in efficiency is caused by time pressure (the leading suspect) and imperfect competition.

In all other conditions, efficiency starts below the level observed in the control condition because commercial practices lead to higher prices and protective measures only partially counteract this effect. However, after the first round, efficiency increases faster than in the control condition, eventually reaching approximately the same level (57% for commercial practices, 48% for formal sanctions, 55% for informal sanctions and 56% for regret). If we delete round five ('end of the game effect'), only the treatment with formal sanctions reaches a similar level of efficiency to the control condition (59% vs 57%). Thus, the alternative hypothesis in relation to H4 is supported.

On average, buyers (mean 42.53, SD 22.28) gained more than sellers (mean 33.40, SD 12.75) but with greater variability.

Figure 3 shows the quantity exchanged (panel a), the average transaction price (panel b), gross profits (panel c) and net profits (panel d) across rounds and by

---

[6]It is important to recall that prices were paid out of the valuation of the package, as the buyers were gaining the difference between the two. This is standard in market experiments.
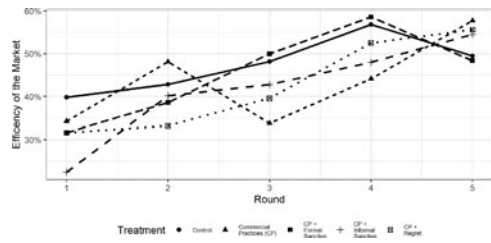
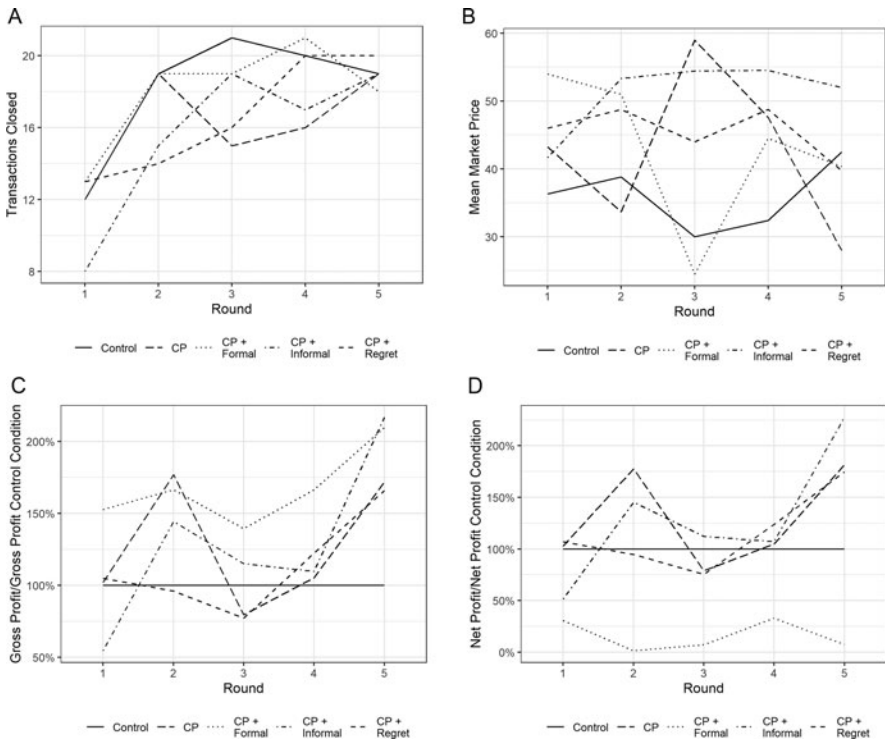**Figure 2.** Efficiency by experimental condition

Figure 3. The dynamics of (a) quantity, (b) price, (c) gross profits and (d) net profits

treatment. It can be seen from panel a that the average quantity in all treatments is lower than that in the control, and the difference increases when informal sanctions are used, driven by a lower willingness to buy. Panel b shows that prices remain higher under commercial practices, as per H1. According to panels c and d, sellers' earnings in the presence of commercial practices are higher than those in the control (the result is clear in all treatments where commercial prices are present). Nevertheless, part of the sellers' rents comes from cheating, and sanctions constrain this source of income. Supplementary Table A1 shows the differences across treatments for the four outcome variables, prices of sealed transactions, likelihood to seal a transaction, gross profits and net profits, estimated using OLS.

Discussion and conclusions

The results of this study have several implications. When sellers have access to commercial practices, they try to extract rents by cheating. By providing misleading information, they try to deceive consumers, possibly relying on the time pressures faced by buyers. The result of this cheating, in the form of higher prices, reduces social welfare.

However, the introduction of protective measures affects the final market outcome. We found that the possibility of formal sanctions changed the types of practices that were adopted, reducing the use of BPG and increasing the use of DP and RP. Similar changes occurred in the cases of informal sanctions and regret. In all cases, the net effect was zero. This change is an endogenous response to the incentive to cheat. Formal sanctions reduce cheating and induce sellers to shift from BPG to DP and RP. Informal sanctions and regret generate a similar adjustment, but with zero net effect. Thus, the policy implication is that blacklisting a specific practice might have a limited effect on aggregate cheating because other practices will serve the same purpose.

Access to commercial practices matters in terms of aggregate efficiency. It drives prices upwards and seems to change the balance of bargaining power between sellers and buyers. Auditing is effective in terms of mitigating these effects, whereas the impact of informal sanctions is less predictable. In fact, on the one hand, cheating is mitigated by negative incentives. On the other hand, buyers can engage in anti-social behaviour by reporting non-cheaters. An unintended consequence of this informal sanction regime is that consumers become more vigilant and less likely to post an order. The regret nudge was ineffective.

A study is usually judged in terms of both internal validity – to what extent a fact or mechanism has been correctly identified by an empirical study – and external validity – to what extent it can be generalized (Guala, 2005). First, we used the gold standard of induced value – making transactions incentive-compatible – and random assignment of treatments to eliminate selection bias. This should ensure an high level of control and internal validity. Second, we collected data in two countries and used an ecologically valid setting involving framed goods and time pressure, similar to a real-world platform. These are elements of external validity of the study. It is possible that other elements of the study are less likely to generalize. For instance, buyers and sellers in the lab behave as average consumers but in real markets, where sorting takes place, prices and quantities are determined by marginal consumers and producers (List, 2002). We take a cautious stance: on the one hand, it is true that real sellers are companies and not individuals, thus following a different decision-making process. On the other hand, markets shape behaviour in ways that is consistent with predictions from theories of incentives (Smith, 1982). Further evidence is needed for this class of market experiments.

This study also has some limitations. We only examined three commercial practices, and while the experimental market proved to be manageable, future studies should assess the impact of other practices including controversial practices such as data profiling and dark patterns (Mathur *et al.*, 2019).

As an extension of this work, it would be worthwhile comparing different levels of vulnerability. In Europe and the US, the regulator distinguishes between the average consumer – attentive and circumspect – and the vulnerable consumer. In this study,

consumers were under time pressure because of the first-come-first-served rule, but emotional vulnerabilities can also play a role in purchase decisions. A further extension would be to consider different clearing institutions, such as the double auction, in an attempt to identify the effects of interactions with various protective measures.

References

Allcott, H. and S. Mullainathan (2010), 'Behavior and energy policy', *Science*, **327**(5970): 1204–1205. https://doi.org/10.1126/science.1180775.

Baghestanian, S., V. Lugovskyy, D. Puzzello and S. Tucker (2014), 'Trading Institutions in Experimental Asset Markets: Theory and Evidence', *Working Paper,* http://dx.doi.org/10.2139/ssrn.3968193.

Battigalli, P. and M. Dufwenberg (2007), 'Guilt in games', *American Economic Review*. https://doi.org/10.1257/aer.97.2.170.

Becker, G. S. (1968), 'Crime and Punishment: an Economic Approach', in *The Economic Dimensions of Crime*, UK: Palgrave Macmillan, 13–68. https://doi.org/10.1007/978-1-349-62853-7_2.

Bicchieri, C. (2006), *The Grammar of Society: The Nature and Dynamics of Social Norms.* Cambridge: Cambridge University Press.

Bordalo, P., N. Gennaioli and A. Shleifer (2013), 'Salience and Consumer Choice Nicola Gennaioli Andrei Shleifer', *Journal of Political Economy*, **121**(5): 803–843.

Calo, R. (2014), 'Digital market manipulation', *George Washington Law Review*, **82**(4): 995–1051. https://doi.org/10.2139/ssrn.2309703.

Celse, J., S. Max, W. Steinel, I. Soraperra and S. Shalvi (2019), 'Uncertain lies: How payoff uncertainty affects dishonesty', *Journal of Economic Psychology*. https://doi.org/10.1016/j.joep.2018.09.003.

Chamberlin, E. H. (1948), 'An experimental imperfect market', *Journal of Political Economy*, **56**(2): 95–108. https://doi.org/10.1086/256654.

Charness, G. and M. Dufwenberg (2006), 'Promises and partnership', *Econometrica*, **74**(6): 1579–1601.

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. https://doi.org/10.1016/j.jbef.2015.12.001.

Coleman, J. S. (1990), *Foundations of Social Theory.* Cambridge: The Belknap Press of Harvard University Press.

Coricelli, G., H. D. Critchley, M. Joffily, J. P. O'Doherty, A. Sirigu and R. J. Dolan (2005), 'Regret and its avoidance: A neuroimaging study of choice behavior', *Nature Neuroscience*. https://doi.org/10.1038/nn1514.

Coricelli, G., E. Rusconi and M. C. Villeval (2014), 'Tax evasion and emotions: An empirical test of re-integrative shaming theory', *Journal of Economic Psychology*. https://doi.org/10.1016/j.joep.2012.12.002.

Crosetto, P. and A. Gaudeul (2017), 'Choosing not to compete: Can firms maintain high prices by confusing consumers?', *Journal of Economics & Management Strategy*, **26**(4): 897–922. https://doi.org/10.1111/jems.12212.

Davis, D. and C. A. Holt (1993), *Experimental Economics*. Princeton: Princeton University Press.

Dworking, G. (1988), *The Theory and Practice of Autonomy*. Cambridge: Cambridge University Press.

Ellingsen, T., M. Johannesson, S. Tjøtta and G. Torsvik (2010), 'Testing guilt aversion', *Games and Economic Behavior*. https://doi.org/10.1016/j.geb.2009.04.021.

European Parliament and Council (2019), *Directive (EU) 2019/2161 of the European Parliament and of the Council of 27 November 2019 Amending Council Directive 93/13/EEC and Directives 98/6/EC, 2005/29/EC and 2011/83/EU of the European Parliament and of the Council as regards the Better Enforcement and Modernisation of Union Consumer Protection Rules*.

Gneezy, U. and A. Kajackaite (2020), 'Externalities, stakes, and lying', *Journal of Economic Behavior and Organization*, **178**: 629–643. https://doi.org/10.1016/j.jebo.2020.08.020.

Gneezy, U. and A. Rustichini (2000), 'A fine is a price', *Journal of Legal Studies*, **XXIX**(January): 1–17.

Gu, Y. and T. Wenzel (2015), 'Putting on a tight leash and levelling playing field: An experiment in strategic obfuscation and consumer protection', *International Journal of Industrial Organization*, **42**: 120–128. https://doi.org/10.1016/j.ijindorg.2015.07.008.

Gu, Y. and T. Wenzel (2020), 'Curbing obfuscation: Empower consumers or regulate firms?', *International Journal of Industrial Organization*, **70**: 102582. https://doi.org/10.1016/j.ijindorg.2020.102582.

Guala, F. (2005), *The Methodology of Experimental Economics*. Cambridge: Cambridge University Press.

Hanson, J. and D. Kysar (1999), 'Taking behavioralism seriously: Some evidence of market manipulation', *Harvard Law Review*, **112**(7): 1420–1572.

Harrison, G. W. and J. A. List (2004), 'Field experiments', *Journal of Economic Literature*, **42**(4): 1009–1055.

Hermann, B., C. Thoni and S. Gachter (2008), 'Antisocial punishment across societies', *Science*, **319**: 1362–1367.

Homans, G. C. (1961), *Social Behavior: Its Elementary Forms*. New York: Harcourt Brace & World.

Huck, S. and B. Wallace (2015), The Impact of Price Frames on Consumer Decision-making Experimental Evidence. Working Paper.

Irlenbusch, B., T. Mussweiler, D. J. Saxler, S. Shalvi and A. Weiss (2020), 'Similarity increases collaborative cheating', *Journal of Economic Behavior and Organization*. https://doi.org/10.1016/j.jebo.2020.06.022.

Jabłonowska, A., M. Kuziemski, A. M. Nowak, H.-W. Micklitz, P. Pałka and G. Sartor (2018), 'Consumer law and artificial intelligence: Challenges to the EU consumer law and policy stemming from the business' use of artificial intelligence. Final report of the ARTSY project', *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3228051.

Kalaycı, K. (2016), 'Confusopoly: competition and obfuscation in markets', *Experimental Economics*, **19**(2): 299–316. https://doi.org/10.1007/s10683-015-9438-z.

Kalayci, K. and J. Potters (2011), 'Buyer confusion and market prices', *International Journal of Industrial Organization*, **29**(1): 14–22. https://doi.org/10.1016/j.ijindorg.2010.06.004.

Keller, B. P. (1996), '"It keeps going and going and going": The expansion of false advertising litigation under the Lanham Act', *Law and Contemporary Problems*, **59**(2): 131. https://doi.org/10.2307/1192073.

Lewis, M. (1992), *Shame: The Exposed Self*. New York: Free Press.

List, J. (2002), 'Testing neoclassical competitive market theory in the field', *Proceedings of the National Academy of Sciences*, **99**(24): 15827–15830.

Loomes, G. and R. Sugden (1982), 'Regret theory: An alternative theory of rational choice under uncertainty', *The Economic Journal*. https://doi.org/10.2307/2232669.

Lopez, M. C., J. J. Murphy, J. M. Spraggon and J. K. Stranlund (2012), 'Comparing the effectiveness of regulation and pro-social emotions to enhance cooperation: Experimental evidence from fishing communities in Colombia', *Economic Inquiry*, **50**(1): 131–142. https://doi.org/10.1111/j.1465-7295.2010.00344.x.

Luguri, J. and L. J. Strahilevitz (2021), 'Shining a light on dark patterns', *Journal of Legal Analysis*, **13**(1): 43–109. https://doi.org/10.1093/jla/laaa006.

Luguri, J. and L. J. Strahilevitz (2021), 'Shining a light on dark patterns', *Journal of Legal Analysis*, **13**(1): 43–109.

Masclet, D., C. Noussair, S. Tucker and M. C. Villeval (2003), 'Monetary and nonmonetary punishment in the voluntary contributions mechanism', *American Economic Review*. https://doi.org/10.1257/000282803321455359.

Mathur, A., G. Acar, M. J. Friedman, E. Lucherini, J. Mayer, M. Chetty and A. Narayanan (2019), 'Dark patterns at scale: Findings from a crawl of 11 K shopping websites', *Proceedings of the ACM on Human-Computer Interaction*, **3**(CSCW): 1–32. https://doi.org/10.1145/3359183.

Mehta, J., C. Starmer and R. Sugden (1994), 'The nature of salience: An experimental investigation of pure coordination games', *American Economic Review*, **84**(3): 658–673.

Netherlands Authority for Consumers & Markets (2020), *Protection of the online consumer: Boundaries of Online Persuasion*. The Hague: ACM.

Normann, H. T. and T. Wenzel (2019), 'Shrouding add-on information: An experimental study', *Scandinavian Journal of Economics*, **121**(4): 1705–1727. https://doi.org/10.1111/sjoe.12319.

Rasch, A., M. Thöne and T. Wenzel (2020), 'Drip pricing and its regulation: Experimental evidence', *Journal of Economic Behavior and Organization*, **176**: 353–370. https://doi.org/10.1016/j.jebo.2020.04.007.

Richards, N. and W. Hartzog (2019), 'The pathologies of digital consent', *Washington University Law Review*, **96**: 1461.

Rick, S. and G. Loewenstein (2008), 'The Role of Emotions in Economic Behavior', in M. Lewis, J. Haviland-Jones, and L. Feldman (eds), *Handbook of Emotions*, New York: The Guilford Press.

Sahni, N. S. and H. S. Nair (2020), 'Sponsorship disclosure and consumer deception: Experimental evidence from native advertising in mobile search', *Marketing Science*, **39**(1): 5–32. https://doi.org/10.1287/mksc.2018.1125.

Smith, V. L. (1962), 'An experimental study of competitive market behavior', *Journal of Political Economy*, **70**(2): 111–137. https://doi.org/10.1086/258609.

Smith, V. L. (1982), 'Microeconomic systems as an experimental science', *The American Economic Review*, **72**(5): 923–955. https://doi.org/10.1017/cbo9780511528354.018.

Stuyck, J. (2015), 'The Court of Justice and the Unfair Commercial Practices Directive', *Common Market Law Review*, **52**(3): 721–752.

Thaler, R. (1980), 'Toward a positive theory of consumer choice', *Journal of Economic Behavior and Organization*, **1**(1): 39–60. https://doi.org/10.1016/0167-2681(80)90051-7.

Thaler, R. H. and C. R. Sunstein (2003), 'Libertarian paternalism', *American Economic Review*, **93**(2): 175–179. https://doi.org/10.1257/000282803321947001.

Thaler, R. H. and C. Sunstein (2008), *Nudge*. New Haven and London: Yale University Press.

van de Ven, J. and M. C. Villeval (2015), 'Dishonesty under scrutiny', *Journal of the Economic Science Association*, **1**(1): 86–99. https://doi.org/10.1007/s40881-015-0002-6.