THE
ECONOMIC
JOURNAL

The Economic Journal, **133** (*May*), 1430–1465 https://doi.org/10.1093/ej/ueac093
Advance Access Publication Date: 15 February 2023

STEERING FALLIBLE CONSUMERS*

Paul Heidhues, Mats Köster and Botond Kőszegi

Online intermediaries with information about a consumer's tendencies often 'steer' her toward products she is more likely to purchase. We analyse the welfare implications of this practice for 'fallible' consumers, who make statistical and strategic mistakes in evaluating offers. The welfare effects depend on the nature and quality of the intermediary's information and on properties of the consumer's mistakes. In particular, steering based on high-quality information about the consumer's mistakes is typically harmful, sometimes extremely so. We argue that much real-life steering is of this type, raising the scope for a broader regulation of steering practices.

Internet companies like Facebook or Google harness their tremendous knowledge about users' tendencies to help sellers make sales. One primary way they do so is through *steering*—influencing which products a consumer considers for purchase (e.g., Agarwal *et al.*, 2016; Eliaz and Spiegler, 2016; Crémer *et al.*, 2019; de Corniere and Taylor, 2019; CMA, 2020; Monopolkommission, 2020). Existing research on this phenomenon (e.g., Varian, 1996; Bergemann and Bonatti, 2011; Hagiu and Jullien, 2011; Inderst and Ottaviani, 2012a; de Corniere, 2016; Marotta *et al.*, 2018; ACCC, 2019; Furman *et al.*, 2019; CMA, 2020; Hidir and Vellodi, 2021; Teh and Wright, 2022) typically assumes that consumers are rational, and steering is based on information about their preferences. In this paper, we analyse the welfare implications of steering by modifying both of these assumptions. We allow for 'fallible' consumers, who make statistical and strategic mistakes in evaluating purchase options. And accordingly, we allow for steering based on information about a consumer's mistakes rather than her true preferences. The implications depend on the type of information used for steering and the type of mistakes the consumer makes. But in many arguably relevant situations, steering lowers consumer welfare, sometimes drastically so. Hence, there is a case for regulating the steering practices of internet companies.

We introduce our formal framework in Section 2. An intermediary offers exactly one of I ex ante identical products to a consumer for purchase; call this product i^*. The consumer observes a signal $w_{i^*} = v_{i^*} + m_{i^*}$, where v_{i^*} is her true value from the product and m_{i^*} is a mistake in assessing this value. The consumer's subjective expected value is $\tilde{v}_{i^*} = \tilde{v}(w_{i^*})$ for a fixed increasing function $\tilde{v}(\cdot)$ that does not depend on the intermediary's behaviour. The price of each good is fixed and normalised to zero, so the consumer buys if and only if $\tilde{v}_{i^*} \geq 0$.

The intermediary, in turn, understands the consumer's behaviour, and aims to maximise the probability of purchase. In doing so, it uses information about the consumer to select i^*. Under value-based steering, it uses information about the values v_i. It may, for example, recommend

* Corresponding author: Paul Heidhues, Düsseldorf Institute for Competition Economics (DICE), Heinrich-Heine University, Universitätsstraße 1, 40225 Düsseldorf, Germany. Email: heidhues@dice.hhu.de

This paper was received on 12 July 2021 and accepted on 16 December 2022. The Editor was Amanda Friedenberg.

We are grateful to Heski Bar-Isaac, Amanda Friedenberg, Justus Haucap, Johannes Johnen, Martin Peitz, Heidi Thysen and three anonymous referees for helpful feedback. Heidhues and Kőszegi thank the European Research Council for financial support under Grant #788918. Köster thanks the German Research Foundation (project number: 235577387/GRK1974) for doctoral funding.

Within the last three years – in collaboration with E.CA Economics – Heidhues has engaged in competition and consumer protection consulting for the Competition and Markets Authority of the UK as well as in the context of trucking, banking, elevator, and timber industries. This paper is neither funded through any of this work nor was any payment for his consulting services conditional on the findings in his academic research.

the medically best treatment for a condition the consumer has been investigating. This is likely truly useful for the consumer (v_{i^*} is high), but what she thinks about it (m_{i^*} or $\tilde{v}_{i^*}$) is unclear. Analogously, under mistake-based steering, the intermediary uses information about the mistakes m_i. It might, for example, recommend a credit card with complex fees to a consumer who displays financial illiteracy in her messages. Such a consumer is unlikely to notice many of the fees ($m_{i^*} = w_{i^*} - v_{i^*}$ is high). Finally, under perceived-value-based steering, the intermediary uses information about the perceived values $\tilde{v}_i$. It might, for example, recommend tickets to a play whose printed version the consumer has purchased previously. Such behaviour indicates that she perceives the play as providing high value ($\tilde{v}_{i^*}$ is high). We compare consumer welfare under steering to that under no steering, in which case the intermediary's recommendation is random.

Our specification allows for two kinds of fallibility. First, the exogeneity of $\tilde{v}(\cdot)$ implies that the consumer does not account for the steering going on; she is strategically naive. Second, the form of $\tilde{v}(\cdot)$ can capture statistical errors the consumer makes even absent steering.

To complete the framework, we introduce notions regarding the consumer's mistakes and the quality of the intermediary's information. While not obvious ex ante, these notions are key in determining the welfare effect of steering. We say that the consumer 'buys reasonably' if a random product's average value conditional on purchase is positive. This means that the consumer benefits from being offered a random product. Analogously, the consumer 'refrains reasonably' if a random product's average value conditional on no purchase is negative. We also say that steering is strong if it is likely to induce purchase. Conversely, steering is weak if the intermediary's information is poor. Then, steering changes the probability of purchase only minimally.

In Section 3, we consider benchmarks with fully rational consumers. A simple case is when—as in most of the literature—the consumer evaluates products with no noise (m_i is degenerate at zero). Then, steering must be value based. Analogously to previous arguments, we observe that steering increases the consumer's welfare by offering her a better selection of products. Even with noise (non-degenerate m_i), a rational consumer benefits from value-based and perceived-value-based steering. And under a wide class of distributional assumptions, she also benefits from strong mistake-based steering. The latter occurs because, after applying a heavy discount to her signal, she can estimate product values accurately.

In Section 4, we analyse value-based steering with fallible consumers. Under weak steering, the consumer is offered a product that is close to random. This lowers her welfare if she does not benefit from being offered a random product, i.e., she does not buy reasonably. Being based on values, however, steering results in a better-than-random offer. This raises the consumer's welfare if she benefits from being offered a random product, i.e., she buys reasonably. Furthermore, strong steering pre-selects largely valuable products, so it is welfare increasing for any consumer.

In Section 5, we turn to mistake-based steering. Again, under weak steering, the consumer is offered a near-random product. This increases her welfare if she benefits from being offered a random product, i.e., she buys reasonably. Being independent of values, however, mistake-based steering raises the probability of purchase without improving product selection for the consumer. Hence, it lowers her welfare if she does not benefit from random product selection, i.e., she does not buy reasonably. Furthermore, if steering is strong then it induces the consumer to buy with near certainty. This means that she buys the products she would have refused absent steering. She benefits if these products are on average valuable, i.e., if she does *not* refrain reasonably.

In Section 6, we analyse perceived-value-based steering. Being partly based on values, steering improves the selection of products to better than random. Hence, just like value-based steering, perceived-value-based steering increases the consumer's welfare if she buys reasonably. In contrast, consider steering that eliminates some products (with low perceived values) the consumer would have rejected anyway. Such steering raises the probability of purchase, but leaves the consumer's value conditional on purchase unchanged. As a result, it hurts her if she does not buy reasonably.

These results allow us to separately identify the effects of strategic naivete and statistical errors on the steered consumer's welfare. If the consumer makes no statistical errors then she both buys and refrains reasonably. Like a rational consumer, therefore, she benefits from value-based and perceived-value-based steering. Unlike a rational consumer, however, she is harmed by strong mistake-based steering. If the consumer makes statistical errors in addition to being strategically naive then all forms of steering can be harmful. At the same time, any type of steering benefits the consumer if she does not refrain reasonably. Intuitively, such a consumer buys too rarely, so she benefits from being induced to buy more often. This raises the theoretical possibility that any mistake-based steering is beneficial.

Although a full analysis is beyond the scope of the paper, in Section 7 we begin to identify the implications of endogenous prices. In particular, we consider steering with many products when the intermediary has accurate information and helps with setting profit-maximising prices. Then, the intermediary is likely to find a product the consumer really wants, and charge a high price. Still, the consumer benefits from value-based steering. Intuitively, the intermediary is concerned that she will mistakenly fail to purchase, so it prices well below her value. For mistake-based and perceived-value-based steering, in contrast, the intermediary prices aggressively to exploit the consumer's mistake. This harms the consumer and often yields large negative consumer welfare.

The above results indicate that the impact of steering is sensitive to details. Nevertheless, in Section 8 we use a range of suggestive evidence and arguments to draw some tentative specific conclusions. First, there is an ever-increasing abundance of products and ease of obtaining information about consumers. This suggests that the case of strong steering is particularly relevant. Second, much steering is best interpreted as mistake based. This includes steering guided by experiments ('A/B tests') on how to market a given product, and that guided by plausible machine-learning algorithms. Third, mistake-based steering is less beneficial for the intermediary if consumers are rational than if they are fallible. The heavy use of such steering therefore suggests that many consumers must be in the latter category. Fourth, firms have an incentive to induce excessive purchases, for instance by hiding parts of a product's price. As a consequence, reasonability in buying (but not reasonability in refraining) is likely often violated. Applying our results to such a combination of circumstances, we find that steering is prone to being harmful. There is therefore a case for regulating existing practices, or restricting the information that steering can be based on. In particular, we argue that steering based solely on the consumer's self-declared interests is more likely to be beneficial. The same is the case for steering based on self-initiated search for a narrow category of products. We conclude in Section 9 by discussing how these insights relate to recent proposals to regulate steering practices.

1. Related Literature

Our paper is related to large or growing literatures on steering, privacy and price discrimination, and naive or inattentive consumers. No previous paper, however, analyses our main question, the welfare effects of different types of steering when consumers make mistakes. Indeed, almost no paper considers the role of mistakes in steering at all. In our setting, there is a stark contrast. Steering a rational consumer tends to be beneficial if prices do not respond much. Furthermore, the harm with endogenous prices is limited (to reducing consumer surplus to zero). But steering a fallible consumer is often harmful whether or not prices are endogenous. In addition, the harm is unbounded. Thinking about steering using a rational model is therefore misleading.

Many papers (e.g., Varian, 1996; Bergemann and Bonatti, 2011; de Corniere, 2016; Marotta *et al.*, 2018; Hidir and Vellodi, 2021; Teh and Wright, 2022) and policy reports (e.g., ACCC, 2019; Furman *et al.*, 2019; CMA, 2020) argue that holding prices fixed, steering benefits consumers. We show that in the case of fallible consumers, this conclusion is incorrect. Instead, existing research has focused on two other main sources of harm from steering. First, prices may change, with this effect typically being complex and ambiguous (e.g., de Corniere, 2016; de Corniere and de Nijs, 2016; Marotta *et al.*, 2018; Teh and Wright, 2022). For much of the paper, we abstract from these ambiguous pricing effects. Second, the intermediary's incentives—and therefore its recommendations—may not be aligned with consumer preferences.[1] In our model, the bias in recommendations derives from consumer mistakes, leading to completely unrelated insights.

A large literature studies the welfare effects of privacy protection (e.g., Stigler, 1980; Posner, 1981; Hermalin and Katz, 2006; MacCarthy, 2010; Fairfield and Engel, 2015; Acemoglu *et al.*, 2022; Bergemann *et al.*, 2022). Firms being able to observe certain characteristics of a consumer is also central to research on third-degree price discrimination. These literatures focus on classical information-theoretic considerations. Due to our focus on consumer mistakes, we derive completely different results.

Like our model, work on naivete-based discrimination (Eliaz and Spiegler, 2006; Heidhues and Kőszegi, 2010; 2017; Johnen, 2020; Bar-Gill, 2021) allows for differential treatment of consumers based on their mistakes. In these papers, consumers consider all available offers, whereas steering amounts to influencing what the consumer considers.

2. Framework

2.1. *Information-Theoretic Preliminaries*

We first define standard concepts regarding signals s that are continuously or discretely distributed conditional on a random variable z.

1. The signal structure satisfies the monotone likelihood ratio property (MLRP) if the conditional density or probability mass function $\psi(s|z)$ satisfies $\psi(s|z)\psi(s'|z') \geq \psi(s|z')\psi(s'|z)$ for any $z > z'$ and $s > s'$. MLRP implies that a higher signal is 'good news': if $s > s'$ then the distribution of z conditional on s (weakly) first-order stochastically dominates the distribution of z conditional on s' (Milgrom, 1981, Proposition 1). Hence, $\mathbb{E}[z|s]$ is increasing in s.

2. The signal structure is informative if there are z and z' such that $\psi(\cdot|z) \neq \psi(\cdot|z')$.

2.2. *Model*

There are I symmetric sellers of ex ante identical products $i = 1, \ldots, I$. The price of each product is fixed and normalised to zero. An intermediary offers exactly one of the products to a consumer, which the consumer can purchase or not. We denote the recommended product by $i^* \in \{1, \ldots, I\}$. The consumer observes a signal $w_{i^*} = v_{i^*} + m_{i^*}$, where v_{i^*} is her value from the product and m_{i^*} is a mistake in assessing that value. The values v_i are drawn according to the

[1] In particular, recommendations can be directed toward sellers that pay higher commissions (Inderst and Ottaviani, 2012a,b; Murooka, 2015), or (in what is called self-preferencing) toward the intermediary's own products (Crémer *et al.*, 2019; de Corniere and Taylor, 2019; CMA, 2020; Monopolkommission, 2020). Similarly, an intermediary that is paid per click might bias recommendations to generate additional traffic (Hagiu and Jullien, 2011).

cumulative distribution function G with density g. The mistakes m_i are drawn according to the cumulative distribution function F with density f. Both g and f are continuously differentiable and log-concave. All v_i and m_i are independent, they have full support on $\mathbb{R}$ and their hazard rates approach infinity when the underlying variable does. In addition, the structure of signals w_i satisfies MLRP.[2]

The consumer's assessment of product i^* is non-strategic and independent of the intermediary's behaviour. In particular, her subjective mean value is $\tilde{v}_{i^*} = \tilde{v}(w_{i^*})$, where $\tilde{v}(\cdot)$ is an exogenously fixed and strictly increasing function with full range on $\mathbb{R}$. Recalling that prices are zero, the consumer buys product i^* if and only if $\tilde{v}_{i^*} \geq 0$.

The intermediary understands the consumer's behaviour, and uses information about her before selecting i^*. The information comes in the form of a signal $s_i \in \mathbb{R}$ for each i, with three types of steering depending on what s_i is about. Under value-based steering, s_i is about v_i: the s_i are identically distributed conditional on v_i, and are independent of each other and of the m_i. Under mistake-based steering, s_i is about m_i: the s_i are identically distributed conditional on m_i and are independent of each other and of the v_i. Under perceived-value-based steering, s_i is about $\tilde{v}_i$: the s_i are identically distributed conditional on $\tilde{v}_i$, and are independent of each other and of the v_i and m_i conditional on $\tilde{v}_i$. In each case, the signal structure satisfies MLRP and is informative.

Two examples of steering technologies are binary and perfect steering. Under binary steering, the intermediary observes whether the parameter in question (v_i, m_i or $\tilde{v}_i$) lies above a given cutoff. Under perfect steering, the intermediary perfectly observes the relevant parameter. When I is large, perfect steering arguably approximates many online markets of the likely near future. Indeed, there is a large abundance of products online, and intermediaries have plenty of information about an individual consumer's tendencies.

The intermediary receives a fixed commission for a sale, so it aims to maximise the probability of purchase.[3] For each type of steering, a higher signal s_i is good news about the respective parameter. Hence, it is also good news about $\tilde{v}_i$ and the resulting probability of purchase.[4] Therefore, it is optimal for the intermediary to recommend a product for which it observed the highest signal. We assume that it does so.

We analyse the effects of the three types of steering on consumer welfare. To do so, we compare the consumer's expected utility (with the expectation taken over v_i, m_i and s_i) to that under no steering, where the intermediary recommends a random product.[5]

2.3. *Comments and Examples*

The above model allows the consumer to be fallible in two ways. First, because $\tilde{v}(\cdot)$ is exogenously fixed, she is strategically naive: her belief $\tilde{v}_{i^*}$ fails to account for how the intermediary chose i^*. Second, the form of $\tilde{v}(\cdot)$ can capture statistical mistakes in evaluating products she may make even absent steering.

[2] Since v_i and m_i are independent, this is equivalent to assuming that the density f of m_i is such that, for any $x > 0$, $f(m)/f(m+x)$ is increasing in m.

[3] In general, profit maximisation does not always imply that the intermediary maximises the probability of purchase. It may, for instance, disproportionately recommend its own products simply because their margins are higher (self-preferencing). Whenever such considerations are uncorrelated with v_i and m_i, they can be thought of as making the intermediary's signal more noisy. To the extent that some consideration is correlated with v_i, m_i or $\tilde{v}_i$, it can be subsumed under the relevant type of steering. For instance, Murooka (2015) showed that deceptive products—which induce mistaken purchases—often pay higher commissions. Then, steering consumers according to commissions amounts to mistake-based steering.

[4] For a proof, see Lemma 5 in the Appendix.

[5] This arises as the limit case of our model where the s_i are uninformative.

One example of a statistical mistake is projection bias (e.g., Loewenstein *et al.*, 2003). Suppose, for instance, that the consumer's signal about a gadget is how exciting it seems at the moment. Projection bias implies that she overestimates the extent to which this excitement will persist. As a result, her perceived value is too sensitive to current conditions.[6] Another potential statistical mistake is overinference from small samples (Rabin and Vayanos, 2010). The consumer may, for instance, overestimate the informativeness of recent performance in evaluating a mutual fund.

Strategic naivete, a version of which is assumed in most models in the behavioural-industrial-organisation literature,[7] can arise for multiple reasons. First, it might be implied by the consumer's statistical mistake. As a case in point, a consumer suffering from extreme projection bias may take her signal at face value ($\tilde{v}(w_i) = w_i$). Then, she deems it unnecessary to think about the intermediary's behaviour. Second, the consumer might totally forget that she should take into account the intermediary's role. Third, she might ignore the fact that the intermediary has information about her that she does not have, and uses that for steering.

2.4. *Reasonability*

We identify two novel notions of consumer behaviour that are central in determining the welfare effects of steering. The consumer buys product i^* if and only if $\tilde{v}_{i^*} = \tilde{v}(v_{i^*} + m_{i^*}) \geq 0$. Because $\tilde{v}(\cdot)$ is strictly increasing with full range, there exists a $\bar{w} \in \mathbb{R}$ such that she buys if and only if $v_{i^*} + m_{i^*} \geq \bar{w}$, or $m_{i^*} \geq \bar{w} - v_{i^*}$. Hence, her surplus from being shown a random product is

$$\int_{-\infty}^{\infty} v[1 - F(\bar{w} - v)]\, dG(v). \tag{1}$$

To simplify the exposition, we assume that parameters are such that (1) is non-zero. By similar logic, the average value of a random product conditional on the consumer not buying it is proportional to

$$\int_{-\infty}^{\infty} v F(\bar{w} - v)\, dG(v). \tag{2}$$

DEFINITION 1. *The consumer* buys reasonably *if (1) is positive. She* refrains reasonably *if (2) is negative. She is* always reasonable *if $\tilde{v}_i(w_i) = \mathbb{E}[v_i|w_i]$ for all w_i.*

All three notions are relaxations of the concept of fully rational consumer behaviour. The consumer buys reasonably if her purchases from a random selection raise her welfare on average. She refrains reasonably if the products she rejects from a random selection would lower her welfare on average. It is easy to verify the following. (1) Neither of these two notions implies the other. (2) Any consumer satisfies at least one of the two notions.[8] Finally, an always reasonable consumer—while strategically naive—accounts for her own noisy perception in a Bayesian way. Such a consumer satisfies both reasonability in buying and reasonability in refraining not only in expectation, but for every purchase decision.

[6] For example, several papers document that temporary weather conditions affect choices for future, long-term consumption (e.g., Conlin *et al.*, 2007; Simonsohn, 2010; Busse *et al.*, 2015; Chang *et al.*, 2018).

[7] See, e.g., Eliaz and Spiegler (2006), Gabaix and Laibson (2006), Heidhues and Kőszegi (2010) and the discussion of this issue in Heidhues and Kőszegi (2018).

[8] Fact (2) holds because the products the consumer purchases from a random selection are on average better than those she rejects. Formally, $\int_{-\infty}^{\infty} v[1 - F(\bar{w} - v)]dG(v) > [1 - F(\bar{w})]\mathbb{E}[v_i]$ and $\int_{-\infty}^{\infty} v F(\bar{w} - v)dG(v) < F(\bar{w})\mathbb{E}[v_i]$. Hence, for a consumer to *not* refrain reasonably, it has to be true that $\mathbb{E}[v_i] > 0$ and, thus, that she buys reasonably. Analogously, a consumer who does *not* buy reasonably must refrain reasonably.

Based on the existing literature, there are compelling reasons to expect violations of reasonability in buying. We discuss one such reason: consumers ignore part of a product's price when deciding whether or not to buy. Such 'hidden prices' are well documented in the context of offline markets (see Heidhues and Kőszegi, 2018 for a review), but might be especially relevant online.[9] As is realistic in a world with many products, we suppose that the average product is not worth buying (i.e., $\mathbb{E}[v_i] < 0$). We take a baseline distribution of mistakes F_0, and consider right shifts of it due to a hidden price of $M > 0$: $F_M(m) = F_0(m - M)$.[10] We get the following result.

PROPOSITION 1. *Fix any G such that $\mathbb{E}[v_i] < 0$ and any F_0 and $\tilde{v}(\cdot)$. If $F = F_M$ for a sufficiently large M, then the consumer refrains reasonably, but does not buy reasonably.*

If hidden prices are significant, the consumer typically overestimates her surplus. This means that she buys even for substantially negative values. More generally, using product design or marketing, firms can potentially exploit many biases to induce purchases of useless products.[11] Hence, reasonability in buying (but not reasonability in refraining) is likely often violated.

2.5. *Weak versus Strong Steering*

Below, we analyse steering technologies that are weak versus strong in achieving the intermediary's objective of increasing the purchase probability. On one side, we derive results of the form 'statement A holds for "sufficiently weak" steering'. This means that there exists a $\bar{c}$ such that statement A holds for binary steering with a cutoff below $\bar{c}$. Loosely, steering is weak if it is binary with a low cutoff. Then, the intermediary's information is poor, eliminating a small share of products and not distinguishing among the rest. This raises the probability of purchase only minimally. On the other side, we also derive results of the form 'statement A holds for "sufficiently strong" steering'. This means that there exists a $\underline{\pi}$ such that statement A holds if the consumer's probability of purchase is above $\underline{\pi}$. Loosely, steering is strong if the probability of purchase is close to 1.[12] Strong steering requires a large I. For perfect steering to be strong, a large I is sufficient.[13] For binary steering, both the cutoff and the probability of finding a product above the cutoff must be high.[14]

[9] For instance, online sellers use 'drip pricing' or other methods of prominently displaying partial price information. They also manipulate consumers into signing up for extra services without much consideration. These and other prevalent manipulative practices are collectively termed 'dark patterns' (Mathur *et al.*, 2019).

[10] To see that this captures hidden prices, let w_i be the difference between v_i and the transparent component of the price. Since all prices are fixed, a larger hidden price must be associated with a lower transparent component. Hence, it must also be associated with a higher perceived surplus, and therefore a higher m_i.

[11] As an example, the consumer might overweigh a subset of a product's relevant aspects. Then, firms will take steps to make sure she overweighs the valuable aspects. First, firms may develop the attribute that consumers overweigh. Second, firms may advertise (and draw attention to) the attribute their products are better at. Overweighing the more valuable attribute is also consistent with recent models on context-dependent choice (Bordalo *et al.*, 2013; 2015; Kőszegi and Szeidl, 2013). Similarly, mistakes could result from the misinterpretation of available information. In this case, firms will tailor the information to make themselves look favourable. They might, for instance, use persuasive, but incorrect 'models' that organise past data well (Schwartzstein and Sunderam, 2021).

[12] Given the definition of strong steering, the natural symmetric notion of weak steering would require that the probability of purchase increases by a small amount. Our notion is more restrictive. It does not allow for that small increase in the probability to purchase to be concentrated among higher parameters. For the less restrictive, symmetric definition, our claims regarding weak steering below do not hold.

[13] To see this, consider value-based steering; the other cases are similar. Under perfect value-based steering, the consumer buys with probability $\int_{-\infty}^{\infty} 1 - G(\bar{w} - m)^I \, dF(m)$, which goes to 1 as $I \to \infty$.

[14] Consider again value-based steering; the other cases are similar. The probability of finding at least one product with a value above v^c is $1 - G(v^c)^I$. For a fixed cutoff v^c, this goes to 1 as $I \to \infty$. The consumer buys such a product with probability $\int_{-\infty}^{\bar{w}-v^c} \{[1 - G(\bar{w} - m)]/1 - G(v^c)\} \, dF(m) + 1 - F(\bar{w} - v^c)$, which goes to 1 as $v^c \to \infty$.

3. Benchmarks: Fully Rational Consumers

3.1. *Consumers Who Know Their Values*

The vast majority of the literature on steering assumes that consumers assess their values perfectly (i.e., F has point mass at 0, and $\tilde{v}(v_i) = v_i$). Then, any steering is value based. Remark 1 confirms previous arguments in our setting.

REMARK 1. Value-based steering strictly increases the consumer's welfare.

Intuitively, steering improves the selection of products the consumer considers (in terms of first-order stochastic dominance), and a rational consumer can only benefit from this.

3.2. *Consumers Who Observe Their Values with Noise*

Going beyond the literature, we analyse rational consumers who observe their values with noise (F is non-degenerate). We make three modifications to the model above. First, the intermediary and the consumer have a common prior. Second, in the case of perceived-value-based steering, the intermediary's signals s_i are about w_i rather than $\tilde{v}_i$.[15] Third, we look for Bayesian Nash equilibria (BNE) in which the intermediary behaves as before, but the consumer is fully rational. This means that the intermediary still recommends a product for which it observed the highest signal. The consumer, however, does not value a product according to $\tilde{v}(\cdot)$. Instead, she correctly accounts for both her noisy signal and the intermediary's behaviour.

PROPOSITION 2 (WELFARE EFFECTS OF STEERING ON RATIONAL CONSUMERS). *Fix G and F.*

(*i*) *Value-based steering benefits the consumer.*
(*ii*) *Perceived-value-based steering benefits the consumer.*
(*iii*) *For sufficiently large I, perfect mistake-based steering benefits the consumer.*

Value-based steering improves the selection of products, out of which one can make better purchases. This must benefit a rational consumer. Relatedly, perceived-value-based steering can be seen as trying to select a product the consumer would herself choose. Once again, a rational consumer can only benefit from this.[16] More surprisingly, in the empirically relevant case of perfect steering with many products, even mistake-based steering is beneficial. For an intuition, suppose for a moment that the distribution of m_i is bounded from above by $\bar{m}$. Then, m_{i^*}—being the largest of many m_i—is almost certainly near $\bar{m}$. Subtracting $\bar{m}$ from her signal w_{i^*}, therefore, the consumer extracts near-perfect information about her value. Since her information without steering is imperfect, steering benefits her. Going further, our assumptions on F approximate such a situation of bounded support. In the limit with many products, the error associated with

[15] Since a rational consumer's values are endogenous to the intermediary's behaviour, we cannot define the intermediary's signals to be about $\tilde{v}_i$. In our main model, the two definitions are equivalent. There, we define s_i over $\tilde{v}_i$ to be most consistent with the term 'perceived-value-based steering'.

[16] Parts (i) and (ii) of Proposition 2 are, in fact, stronger than stated. Suppose that the consumer does not observe the recommended product's identity i^*, and for any intermediary signal s_{i^*}, buys with interior probability. Then, for value-based and perceived-value-based steering, the intermediary recommends the product for which it observed the highest signal in *any* BNE. Parts (i) and (ii) of Proposition 2 therefore do not rely on the assumption that the intermediary recommends the product for which it observed the highest signal.

the suggested product becomes highly predictable. Hence, the consumer's signal becomes highly informative about her value.

4. Value-Based Steering

PROPOSITION 3 (WELFARE EFFECT OF VALUE-BASED STEERING). *Fix G, F and $\tilde{v}(\cdot)$.*

(*i*) *Sufficiently strong value-based steering benefits the consumer.*
(*ii*) *If the consumer does not buy reasonably, sufficiently weak value-based steering harms her.*
(*iii*) *The following statements are equivalent.*
  (a) *The consumer buys reasonably.*
  (b) *Value-based steering benefits the consumer for any signal structure of the intermediary.*
(*iv*) *If the consumer does not refrain reasonably or is always reasonable, value-based steering benefits her.*

Part (i) says that, for any consumer, sufficiently strong steering raises welfare. If steering is strong, the intermediary tends to recommend a product the consumer is unlikely to reject. This must be a very valuable product, which thus benefits the consumer. More subtly, part (ii) says that if the consumer does not buy reasonably then weak steering lowers her welfare. All weak steering does is to filter out products with extremely low values, which the consumer would almost certainly reject anyway. Such elimination of nearly irrelevant products induces additional purchases that are similar to those absent steering. This harms the consumer if her average purchase from a random selection is harmful, i.e., she does not buy reasonably. Hence, a necessary condition for all forms of value-based steering to improve the consumer's welfare is that she buys reasonably. Part (iii) implies that the same is also sufficient. Steering improves the selection of products, leading to additional purchases that are better than those from a random selection. But a consumer who buys reasonably already benefits from being offered a random product. Hence, she benefits from the additional purchases induced by steering as well. Now recall that both an always reasonable consumer and a consumer who does not refrain reasonably buy reasonably. Using part (iii), therefore, value-based steering must benefit such consumers (part (iv)).

Figure 1 illustrates the above results in a numerical example with binary steering, and a consumer who takes her signal at face value ($\tilde{v}(w_i) = w_i$). We assume that $m_i \sim \mathcal{N}(0, 1)$, and look at two possibilities for the distribution of a product's value: $v_i \sim \mathcal{N}(0, 1)$ and $v_i \sim \mathcal{N}(-2, 1)$. It is easy to check that the consumer buys reasonably in the former case, but not in the latter case. We consider the limit case $I = \infty$, where the intermediary finds a product above the cutoff v^c with probability 1. We plot the welfare effect of steering as a function of v^c. Note that the higher v^c, the better the product selection. As illustrated in the left panel, value-based steering raises the welfare of a consumer who buys reasonably. But as illustrated in the right panel, weak steering is harmful, while strong steering is beneficial for a consumer who does not buy reasonably. Furthermore, the effect of improving product selection is not monotonic, with the consumer being worst off in the intermediate range. Such steering leads to suggestions that are good enough for the consumer to purchase with a decent likelihood, but not good enough to benefit her.

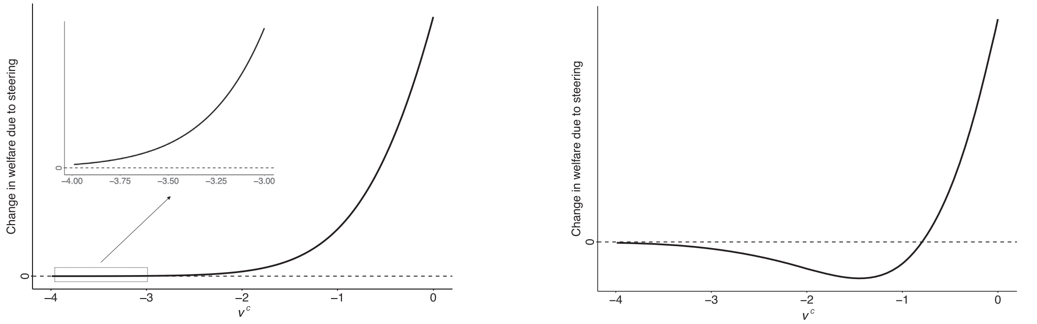

Fig. 1. *Welfare Effect of Binary Value-Based Steering as a Function of the Threshold v^c when $\tilde{v}(w_i) = w_i$, $I = \infty$, $m_i \sim \mathcal{N}(0, 1)$ and $v_i \sim \mathcal{N}(0, 1)$ (Left) or $v_i \sim \mathcal{N}(-2, 1)$ (Right).*
Notes: For visual clarity, the part of the left figure where the curve is close to zero is shown enlarged in the inset.

5. Mistake-Based Steering

PROPOSITION 4 (WELFARE EFFECT OF MISTAKE-BASED STEERING). *Fix G, F and $\tilde{v}(\cdot)$.*

(*i*) *The following statements are equivalent.*
 (a) *The consumer does not buy reasonably.*
 (b) *Mistake-based steering harms the consumer for any signal structure of the intermediary.*
(*ii*) *If the consumer buys reasonably, sufficiently weak mistake-based steering benefits her.*
(*iii*) *The following statements are equivalent.*
 (a) *The consumer does not refrain reasonably.*
 (b) *Mistake-based steering benefits the consumer for any signal structure of the intermediary.*
(*iv*) *If the consumer refrains reasonably or is always reasonable, sufficiently strong mistake-based steering harms her.*

The if-and-only-if statements in Proposition 4 have several implications. By part (i), if the consumer does not buy reasonably then mistake-based steering unambiguously harms her. A consumer who does not buy reasonably is already harmed by her choices from a random selection. To make matters worse, steering induces higher mistakes, leading to additional purchases that have even lower values. Hence, these extra purchases must be harmful on average. Part (i) also implies, however, that if the consumer does buy reasonably then mistake-based steering may benefit her. This actually follows from part (ii): sufficiently weak mistake-based steering benefits her in that case. As with value-based steering, the additional purchases induced by weak mistake-based steering are similar to those from a random selection. These benefit the consumer roughly if and only if she buys reasonably.

Part (iii) implies that if the consumer does not refrain reasonably then mistake-based steering benefits her. Furthermore, this sufficient condition is tight: if the consumer refrains reasonably then certain forms of mistake-based steering harm her. Part (iv), in particular, says that strong steering harms her in this case. By implication, strong steering must harm an always reasonable consumer. These results are best understood starting from part (iv). Under strong steering, the

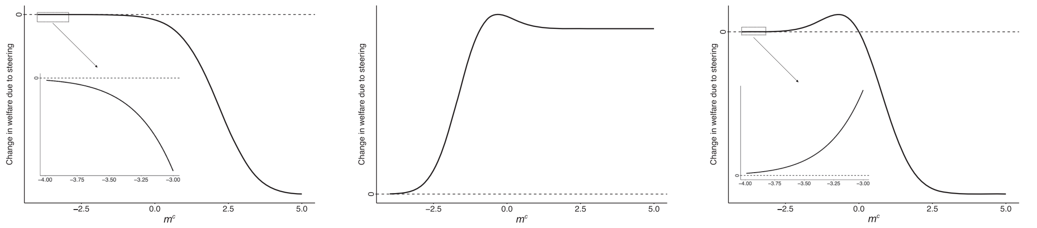

Fig. 2. *Welfare Effect of Binary Mistake-Based Steering as a Function of the Threshold m^c when $\tilde{v}(w_i) = w_i$, $I = \infty$, $m_i \sim \mathcal{N}(0, 1)$, and $v_i \sim \mathcal{N}(-2, 1)$ (Left) or $v_i \sim \mathcal{N}(2, 1)$ (Middle) or $v_i \sim \mathcal{N}(0, 1)$ (Right).*

Notes: For visual clarity, the parts of the left and right figures where the curves are close to zero are shown enlarged in the insets.

consumer almost never refrains. By definition, this is harmful roughly if and only if she would otherwise have refrained reasonably. Going further, extra purchases are less likely to be low value under other forms of steering than under strong steering. Hence, if the consumer benefits from strong steering, she benefits from other forms of steering as well. This is the case if she does not refrain reasonably.

A potential example for part (iii) is life insurance. Evidence suggests that working-age individuals purchase too little life insurance (Gottlieb, 2012), i.e., many do not refrain reasonably. Furthermore, some life-insurance products are designed to be more attractive, albeit economically not more valuable, to consumers (Gottlieb, 2012; Anagol *et al.*, 2017). Indeed, brokers steer consumers toward such products (Anagol *et al.*, 2017). This mistake-based steering can be welfare improving.

The above results have notable implications for a consumer who buys as well as refrains reasonably. Weak steering always benefits such a consumer, but strong steering always hurts her. Intuitively, weak mistake-based steering filters out products that the consumer undervalues, preventing her from refusing valuable products. But strong mistake-based steering identifies products that the consumer overvalues, pushing her toward buying harmful products. This pattern is in contrast with the case of value-based steering. There, strong steering is always welfare increasing, and weak steering can be welfare decreasing.

Figure 2 illustrates the above results by modifying our previous numerical example. We now consider three possibilities for G: $v_i \sim \mathcal{N}(-2, 1)$, in which case the consumer does not buy reasonably; $v_i \sim \mathcal{N}(2, 1)$, in which case she does not refrain reasonably; and $v_i \sim \mathcal{N}(0, 1)$, in which case she both buys and refrains reasonably. We plot the welfare effect of binary steering as a function of the cutoff m^c. Steering always harms a consumer who does not buy reasonably (left panel). Steering always benefits a consumer who does not refrain reasonably (middle panel). And a consumer who buys and refrains reasonably benefits from weak, but not strong steering (right panel).

6. Perceived-Value-Based Steering

PROPOSITION 5 (WELFARE EFFECT OF PERCEIVED-VALUE-BASED STEERING). *Fix G, F and $\tilde{v}(\cdot)$.*

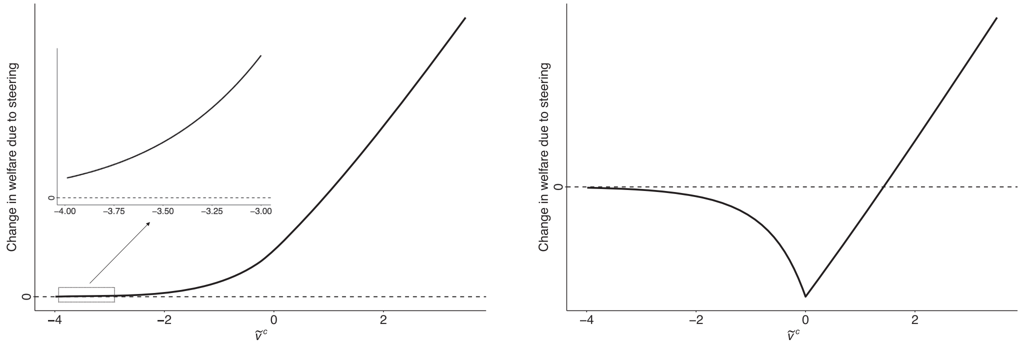

Fig. 3. *Welfare Effect of Binary Perceived-Value-Based Steering as a Function of* $\tilde{v}^c$ *When* $\tilde{v}(w_i) = w_i$, $I = \infty$, $m_i \sim \mathcal{N}(0, 1)$ *and* $v_i \sim \mathcal{N}(0, 1)$ *(Left) or* $v_i \sim \mathcal{N}(-2, 1)$ *(Right).*
Notes: For visual clarity, the part of the left figure where the curve is close to zero is shown enlarged in the inset.

(i) *The following statements are equivalent.*
(a) *The consumer buys reasonably.*
(b) *Perceived-value-based steering benefits the consumer for any signal structure of the intermediary.*
(ii) *If the consumer does not refrain reasonably or is always reasonable, perceived-value-based steering benefits her.*

If the consumer buys reasonably then steering always benefits her, but otherwise it can harm her (part (i)). Hence, if she is always reasonable or does not refrain reasonably, she benefits from steering (part (ii)). These patterns are identical to those under value-based steering, but distinct from those under mistake-based steering. As an example, consider binary steering with a cutoff $\tilde{v}^c \leq 0$. This raises the probability of purchase by eliminating from consideration some products that the consumer would not have bought anyway. Hence, it leaves the distribution of v_i conditional on purchase, and therefore also the consumer's expected utility conditional on purchase, unchanged. By definition, the consumer benefits if and only if she buys reasonably. Suppose, in particular, that the consumer does not buy reasonably. Then the higher the cutoff $\tilde{v}^c \leq 0$, the higher the probability of purchase, and hence the lower the consumer's welfare. In this sense, the better the intermediary's technology is in filtering out products the consumer avoids, the worse off she is.

Figure 3 illustrates the above results using the same example with binary steering as in Figure 1. If the consumer buys reasonably (left panel), steering always benefits her. If she does not buy reasonably (right panel), she benefits from strong, but not from weak steering. In the latter case, her welfare is minimised at the cutoff $\tilde{v}^c = 0$, exactly where the probability of purchase reaches 1.

7. Endogenous Prices

Fully analysing the effects of steering with endogenous prices is beyond the scope of this paper. We therefore consider a special case that is likely to be especially relevant in the near future.

Suppose that steering is perfect and that there are many products. Consistent with our previous framework, the intermediary receives a fixed share of sellers' profits. This implies that the intermediary has an interest in maximising the expected profit from a sale. Accordingly, we assume that after choosing the product to suggest, the intermediary chooses the profit-maximising price for the seller.[17] Just like in our previous model, it is then obviously optimal for the intermediary to choose a product with the highest signal.

With many available products, the recommended one is likely to have a high parameter. Hence, we ask whether the consumer benefits from being steered toward such a product. To simplify the analysis, we assume that the perceived valuation $\tilde{v}(w_i)$ is differentiable in the signal w_i. Furthermore, we make a weak assumption on this derivative: for sufficiently high w_i, $\tilde{v}'(w_i) \leq 1$. Intuitively, this means that the consumer either takes an increase in a high w_i at face value, or discounts it.

The welfare effect of steering again depends on its type.

PROPOSITION 6 (WELFARE EFFECT OF PERFECT STEERING WITH MANY PRODUCTS). *Fix any G and F.*

(*i*) *If the selected v_{i^*} is sufficiently high then value-based steering benefits the consumer.*
(*ii*) *If the selected m_{i^*} is sufficiently high then mistake-based steering harms the consumer.*
(*iii*) *Suppose that $\mathbb{E}[\tilde{v}_i - v_i|\tilde{v}_i]$ is weakly increasing, and non-negative for sufficiently high $\tilde{v}_i$. If the selected $\tilde{v}_{i^*}$ is sufficiently high then perceived-value-based steering harms the consumer.*

A concise way to think about these results is in terms of how pricing responds to the consumer's perception. Under mistake-based steering, the intermediary knows that the consumer overvalues the suggested product, and prices aggressively to exploit this mistake. Under perceived-value-based steering, the intermediary extracts all of the consumer's perceived value, which tends to overestimate her true value. Under value-based steering, the intermediary is worried that the consumer makes the mistake of not buying a high-value product. To mitigate this mistake, it prices more carefully. Hence, the consumer is hurt in the former cases, but benefits in the latter case.

We now flesh out this logic. Under value-based steering, the intermediary suggests a product of high value, enabling high margins from a sale. In pursuit of such margins, the intermediary sets a price that induces the consumer to purchase with high probability. This implies that the intermediary must price well below the consumer's value. Hence, the consumer's tendency to make mistakes guarantees her a large positive surplus. Without steering, in contrast, margins are lower, so the profit-maximising price leaves a smaller surplus to the consumer. Our finding that strong value-based steering benefits the consumer is therefore robust to allowing for endogenous prices.

Under mistake-based steering, the intermediary identifies a product that the consumer overvalues by a large margin. This again leads to high prices and large margins, so that the profit-

[17] An equivalent model arises under natural alternative assumptions. This is the case if the intermediary merely recommends the optimal price, which the seller is happy to accept. It is also the case if the intermediary discloses its information, and the seller sets the price. Our assumption that the intermediary helps sellers maximie profits, potentially by assisting them on pricing, is consistent with some existing practices (e.g., AirBnB: https://www.airbnb.com/resources/hosting-homes/a/setting-a-pricing-strategy-15). Furthermore, the intermediary can, for instance, use an auction as in de Corniere and de Nijs (2016) or Marotta *et al.* (2018) to steer consumers to the seller that earns the highest expected profit.

Table 1. *Sufficient Conditions on Reasonability for Different Types of Steering to be Harmful or Beneficial.*

Type/strength		General	Weak	Strong	Strong + end. price
Value based	Harmful	–	¬ BR	Never	Never
	Beneficial	BR	BR	Always	Always
Mistake based	Harmful	¬ BR	¬ BR	RR	Always
	Beneficial	¬ RR	BR	¬ RR	Never
Perceived value based	Harmful	–	¬ BR	–	Always
	Beneficial	BR	BR	BR	Never

Notes: Arguments in Section 8.2 suggest that strong mistake-based and perceived-value based steering—the grey cells in the table—are the most common steering practices in online markets. BR: the consumer buys reasonably. RR: the consumer refrains reasonably.

maximising price is likely to induce purchase. Since v_{i^*} is independent of m_{i^*}, the consumer is therefore likely to purchase an average product at a very high price. Hence, endogenous pricing amplifies the negative welfare effect of mistake-based steering. Unlike with exogenous prices above, here mistake-based steering harms the consumer even if she does not refrain reasonably. In this case, steering is socially beneficial, but harmful to the consumer.

Our analysis for perceived-value-based steering relies on an additional assumption. The assumption says that a higher perceived valuation reflects a higher mean error, and a high perceived valuation reflects a positive mean error. These properties are natural in the context of statistical mistakes. Then, when the intermediary finds a product with a high perceived value, the consumer overvalues this product. Furthermore, since the intermediary sets the price equal to the consumer's perceived value, her welfare is the negative of her overvaluation. Absent steering, in contrast, the consumer is likely to be offered a product of lower perceived value. By our assumption, she overvalues such a product by less. Hence, if she purchases at a price equal to her perceived value, she is already better off than with steering. And if she purchases at a lower price or does not purchase, she is also better off. Steering thus clearly harms the consumer.

8. Discussion

8.1. *Overview*

Table 1 summarises the insights from Propositions 3 through 6, with some noteworthy patterns. An obvious overall observation is that the effect of steering depends on its type. But there are also some overarching properties. First, any type of weak steering is beneficial if and only if the consumer buys reasonably. A second point follows from the fact that a consumer who does not refrain reasonably always buys reasonably. Namely, a sufficient condition for any steering to increase the consumer's welfare is that she does not refrain reasonably.

We can also determine how different forms of naivete affect the welfare impact of steering. First, consider a strategically naive, but always reasonable consumer. Like her rational counterpart analysed in Section 3, this consumer always benefits from value-based and perceived-value-based steering. In contrast to a rational consumer, however, she is always harmed by perfect mistake-based steering with many products. Second, suppose that the consumer also violates a notion of reasonability. Then, some forms of value-based and perceived-value-based steering may harm her (if she does not buy reasonably). In addition, mistake-based steering may be generally beneficial for her (if she does not refrain reasonably).

While we focus on consumer welfare, one can use the same formalism to analyse social welfare. To do so, we define v_i as the total social welfare from purchase. In addition, we incorporate the consumer's failure to internalise the seller's profit into the mistake m_i. Then, the notion of buying reasonably says that the consumer's average purchase is socially beneficial. Analogously, the notion of refraining reasonably means that the average product the consumer rejects is socially harmful. In this case, it is unnatural to consider always reasonable consumers: such consumers would have to act in the socially optimal way in every situation. In particular, the consumer may reject a product for which social welfare is positive due to the seller's profit. Hence, refraining reasonably can be easily violated. With these modifications, the results in Table 1 apply unchanged.

8.2. *Implications for Real-Life Steering Practices*

We now apply our results to draw some tentative conclusions regarding real-life steering. Since there is little evidence on steering practices, however, our conclusions are somewhat speculative.

As a crucial input, we argue that intermediaries' practices likely result in mistake-based and perceived-value-based steering. Internet companies run many experiments called 'A/B tests' on what garners purchase interest in different consumer groups (Kohavi *et al.*, 2020, p. 5). We can think of A/B tests as comparing product-framing pairs, where the framing corresponds to how and under what circumstances the product is presented. A consumer can then be steered based on what has been effective in inducing purchase from similar others. This group could, for instance, be consumers with the same demographics who type similar words in a search engine (see Kohavi *et al.*, 2020, p. 20, for an example on Amazon).

Most straightforwardly, A/B testing can lead to perceived-value-based steering. This happens if the intermediary experiments with different product offers with no particular attention to the frames used (Hannak *et al.*, 2014 document this type of steering by Expedia and Hotels.com).

Probably the most important use of A/B testing, however, is not for comparing different products. It is for testing different ways of selling the *same* product. Bing, for instance, found that shifting information from an ad's text to its title increased annual revenues by over \$100 million. Such experimental variation leaves the product (and the information about it) fixed, so it is orthogonal to the product's true value. Hence, these types of A/B tests are only informative about the mistakes in consumers' evaluations. Accordingly, they can only be used for mistake-based steering. Indeed, Bing's finding affected the order in which ads were recommended to consumers (see Figure 1.1 of Kohavi *et al.*, 2020).

In contrast, it does not appear likely that firms use A/B testing for value-based steering. For that to happen, a consumer's choices would have to reflect her true values. For this, in turn, it would be necessary for the different frames to hold a consumer's mistake constant. But, when a large variety of products is available, a high correlation between mistakes seems unlikely.[18]

Beyond A/B testing, intermediaries utilise machine-learning algorithms for steering (McMahan *et al.*, 2013). The patterns these algorithms use to predict behaviour are not understood (e.g., Wills and Tatar, 2012; Datta *et al.*, 2015; Bashir *et al.*, 2019). Nevertheless, it seems clear that

[18] For example, consider projection bias as a source of mistakes. A high correlation between mistakes might be plausible in a narrow category of products, such as when only swimsuits are available for sale and weather-based projection bias is the only mistake the consumer makes. Then, on a hot day, for instance, she overvalues different swimsuits by similar margins. But the evaluation of other types of products is surely subject to different mistakes: weather-based projection bias generates a mistake for winter coats that is negatively correlated with that for swimsuits, and mistakes that are not weather based are likely uncorrelated.

they can lead to mistake-based steering. An algorithm may learn to direct consumers toward sellers with deceptive and high prices.[19] An algorithm may use momentary weather conditions at the consumer's location to take advantage of projection bias.[20] An algorithm may infer a consumer's gullibility from her search behaviour, and offer her deceptive products.[21] Mistake-based steering is especially plausible because algorithms are trained in part on data from A/B tests.[22] And looking ahead, algorithms are under development that predict a user's intoxication level or emotional state. These variables likely correlate with mistakes and can therefore be used for mistake-based steering.[23] In contrast, it is difficult to imagine how an algorithm would be used for value-based steering.

To add to the above, recall two previous observations. First, strong steering likely describes many online markets of the present or near future. Second, reasonability in refraining is rarely violated, while reasonability in buying can often be. Then, our theory says that steering has a harmful or ambiguous effect on fallible consumers (see the grey part of Table 1). Furthermore, to the extent that steering is mistake based, its effect is unambiguously negative.

But our results also point to plausible steering environments that may benefit consumers, at least if prices do not respond much to steering. A simple example is perceived-value-based steering where reasonability in buying is satisfied, such as with familiar products and transparent prices. This not only benefits consumers, but also appears achievable. Already today, users can manually update the 'interest profiles' of likes and dislikes major platforms build about them.[24] It is thus easy to keep a list of interests the consumer has actively confirmed or added. Relatedly, a consumer may express a specific interest by initiating a search for a well-defined, narrow range of products. These choices arguably express perceived value. When steering is based exclusively on such interests and regulation against hidden prices is in place, therefore, consumers may benefit. For instance, concert recommendations based on the consumer's stated music preferences and searches for music and events may be welfare increasing.[25]

[19] Consistent with this possibility—albeit offline—audit studies in the markets for life insurance (Anagol *et al.*, 2017) and financial investments (Mullainathan *et al.*, 2011) indicate that advisors tend to steer consumers toward more overpriced, but economically not superior products.

[20] Indeed, there are 'weather targeting' apps designed for this purpose. See, e.g., http://www.weatherads.io/facebook-weather-targeting, accessed June 3, 2021.

[21] For example, Google showed ads for fraudulent investments to individuals who searched for 'high-return investment'. As a result of such behaviour, the UK's Financial Conduct Authority investigated how Google can be used by financial fraudsters (see Vincent, 2020).

[22] See, for instance, https://adwords.googleblog.com/2011/10/ads-quality-improvements-rolling-out.html (accessed on May 19, 2021) and the discussion thereof in Kohavi *et al.* (2020, p. 14).

[23] Uber recently filed a patent application for an algorithm that predicts a user's likelihood of being intoxicated (see https://t1p.de/5hhp, accessed on May 18, 2021). Spotify developed a speech-recognition software that is able to make recommendations (of songs or ads) based on a user's likely emotional state (see https://www.bbc.com/news/entertainment-arts-55839655, accessed on May 19, 2021).

[24] See https://support.google.com/accounts/answer/2662856?hl=en and https://support.google.com/accounts/answer/1634057 (Google) or https://www.facebook.com/help/378618582856718 (Facebook), all accessed on January 17, 2023. Making user interest profiles available and editable is part of a self-regulatory initiative by the Digital Advertising Alliance.

[25] By assuming that the consumer evaluates only the one product the intermediary suggests, our model abstracts from consumer search and its associated costs. To the extent that consumers would themselves search among a subset of the products and steering makes this search easier, it can benefit consumers via lower search costs independently of its type. This consideration is likely to be most important for self-initiated search—where the consumer has expressed an interest in purchasing and has started a search process—and for products the consumer has purchased repeatedly in the past, adding another argument in favour of steering in these situations. In many situations, however, consumers appear to behave just like in our model: they do little to no search, often not scrolling beyond the first set of suggested options (Koulayev, 2014; de los Santos, 2018).

To complete our discussion, we consider how our conclusions are modified if some consumers are fully rational. As a start, suppose that prices are fixed. Then, rational consumers always benefit from perceived-value-based steering, and could benefit from mistake-based steering as well (Proposition 2). But there may be too few rational consumers for this benefit to outweigh the harm to fallible consumers. Indeed, firms' heavy investment into mistake-based steering suggests that many fallible consumers are present. This is because consumers who understand firms' behaviour and observe firms' investments discount suggestions that result from mistake-based steering. Hence, if most consumers were of this type, mistake-based steering would have a limited or negative effect on purchases.

Finally, suppose that prices are endogenous. Then, strong mistake-based steering can cause arbitrarily large harm to fallible consumers. In contrast, its benefit to rational consumers (if any) is limited. Furthermore, perfect perceived-value-based steering—which leaves consumers with zero perceived surplus—is harmful for rational as well as fallible consumers. Hence, the empirically relevant types of steering are likely harmful even if only a small share of consumers are fallible.

9. Conclusion

Recent regulatory proposals to protect consumers by restricting platforms' 'recommender systems' have been prepared with little guidance from economic theory. Our framework provides economic foundations for some of the proposals. We also hope that it can be used to sharpen specific articles, as well as lead to new proposals in the future.

Our results imply that restricting mistake-based steering is probably beneficial. A number of proposed steps can be seen as directly or indirectly doing so. Fletcher *et al.* (2023) suggest that interface designs should not misdirect consumers and their choice architecture should be neutral. They also propose that 'personalized rankings and targeting [should not] be based on characteristics designed to predict vulnerability'. And the proposal for the Digital Service Act (European Commission, 2020, Article 26) requires platforms to determine the risk of their recommender systems for negative effects.

In addition, restricting steering to be based on self-initiated search and self-declared interests would be beneficial. Alternatively, regulators might direct steering toward value-based steering. A challenge for the former type of regulation is enforcing that no other information is used by an intermediary. But there are methods for making this determination that can be deployed and improved.[26] An impediment to the latter type of regulation is that we do not currently know what practices would result in value-based steering. But the proposal by the European Commission (2020) requires platforms to keep and make available data on their practices, which should help researchers to gain a better understanding.[27] We are sceptical, however, regarding the proposal's requirement that a platform has to allow users to opt out of being steered. While some consumers may like this option for other reasons, a fallible consumer with some experience would probably not. Such a consumer may notice that her perceived values are higher under steering than under no steering. Hence, she thinks that steering is useful in identifying products she enjoys.

[26] By the means of 'controlled browsing' experiments (e.g., varying the use of search terms or stated interests), Wills and Tatar (2012) documented that—in violation of their own policy (https://support.google.com/accounts/answer/1634057, accessed on June 11, 2021)—Google used sensitive information on sexual orientation or health status for targeting ads. Hannak *et al.* (2014) used similar experiments to show little personalised pricing in the United States.

[27] Article 29 requires these platforms to make the main parameters on which recommender systems are based public. Article 30 forces the platforms to keep data on advertising as well as anonymous targeting information, and Article 31 requires the platforms to grant researchers access to data on their targeting behaviour.

Appendix A.

A.1. Restating the Welfare under Steering

We express consumer welfare under steering by conditioning on the event that product i is recommended. Consider value-based steering. When being offered product i, the (strategically naive) consumer buys if and only if $\tilde{v}_i = \tilde{v}(v_i + m_i) \geq 0$. Because $\tilde{v}(\cdot)$ is strictly increasing with full range, there exists a $\bar{w} \in \mathbb{R}$ such that she buys if and only if $m_i \geq \bar{w} - v_i$. And since value-based steering leaves the distribution of the mistake F unchanged, the welfare under steering can be written as

$$\sum_{i=1}^{I} \underbrace{\mathbb{P}[\text{product } i \text{ is recommended}]}_{=1/I \text{ since products are symmetric}} \int v[1 - F(\bar{w} - v)]\, dG(v|i)$$
$$= \int v[1 - F(\bar{w} - v)]\, dG(v|i), \qquad \text{(A1)}$$

where $G(\cdot|i)$ is the cumulative distribution function of v_i conditional on product i being recommended. Let $g(\cdot|i)$ be the corresponding density, and note that $g(\cdot|i) = g(\cdot|j)$ for any i and j.

For a random product, i, G, F and $\tilde{v}(\cdot)$ induce a distribution of perceived values $\tilde{v}_i$. Denote the cumulative distribution function of $\tilde{v}_i$ by H and its density by h. Below, we express the welfare under mistake-based and perceived-value-based steering by defining $F(\cdot|i)$ and $H(\cdot|i)$, with corresponding densities $f(\cdot|i)$ and $h(\cdot|i)$, analogously to $G(\cdot|i)$ above.

A.2. Preliminaries on the Implications of MLRP

We collect well-known results on the implications of the monotone likelihood ratio property. Consider CDFs Φ and Ψ, with densities ϕ and ψ, respectively, and let Φ and Ψ satisfy MLRP.

LEMMA 1. *There exists some $x^* \in \mathbb{R}$ such that $\phi(x) < \psi(x)$ for $x > x^*$ and $\phi(x) > \psi(x)$ for $x > x^*$.*

LEMMA 2. *For any $x \in \mathbb{R}$,*

$$(a) \quad \frac{\Phi(x)}{\Psi(x)} \leq \frac{\phi(x)}{\psi(x)} \leq \frac{1 - \Phi(x)}{1 - \Psi(x)} \qquad \textit{and} \qquad (b) \quad \Psi(x) \geq \Phi(x).$$

Moreover, there exists some $\epsilon > 0$ such that the first inequality in (a) is strict on $(x^*, x^* + \epsilon)$ and the inequality in (b) is strict on $(-\infty, x^* + \epsilon)$. The second inequality in (a) is strict on $(-\infty, x^*)$.

LEMMA 3. *For any $x \in \mathbb{R}$,*

$$(a) \quad \frac{\partial}{\partial x}\left[\frac{\Psi(-x) - \Phi(-x)}{1 - \Psi(-x)}\right] \leq 0 \qquad \textit{and} \qquad (b) \quad \frac{\partial}{\partial x}\left[\frac{\Psi(-x) - \Phi(-x)}{\Psi(-x)}\right] \geq 0.$$

The inequality in (a) is strict on $(-x^*, \infty)$ and, for $\epsilon > 0$, the inequality in (b) is strict on $(-(x^* + \epsilon), -x^*)$.

A.3. Preliminaries on Steering Technologies

Notation. Consider the random variable x_i, which could be either v_i, m_i or $\tilde{v}_i$. Let $\Psi(\cdot|x_i = x)$ be the CDF of the signal s_i that the intermediary observes conditional on $x_i = x$, and let $\psi(\cdot|x_i = x)$ be the corresponding density or probability mass function. Denote by $\psi(\cdot)$ the unconditional density or probability mass function of the intermediary's signal. Similarly, let $\Phi(\cdot|s_i = s)$ be the CDF of x_i conditional on signal $s_i = s$, and let $\phi(\cdot|s_i = s)$ be the corresponding density or probability mass function. Denote by $\phi(\cdot)$ the unconditional density or probability mass function of x_i.

LEMMA 4 (SPECIFIC TECHNOLOGIES). *Perfect and binary steering satisfy MLRP and are informative.*

PROOF. Consider perfect steering; that is, $s_i = x_i$ with probability 1. This signal structure satisfies MLRP: for any $x, x' \in \mathbb{R}$ with $x > x'$ and any $s, s' \in \mathbb{R}$ with $s > s'$,

$$\psi(s_i = s'|x_i = x)\psi(s_i = s|x_i = x') = 0,$$

because either $s > s' \geq x > x'$ and, therefore, $\psi(s_i = s|x_i = x') = 0$ or $x > s'$ and, thus, $\psi(s_i = s'|x_i = x) = 0$. Hence, the signal structure satisfies the definition of MLRP,

$$\psi(s_i = s|x_i = x)\psi(s_i = s'|x_i = x') - \underbrace{\psi(s_i = s'|x_i = x)\psi(s_i = s|x_i = x')}_{=0} \geq 0.$$

Moreover, for any $x \neq x'$, $\psi(\cdot|x_i = x) \neq \psi(\cdot|x_i = x')$; i.e., perfect steering is informative.

Consider binary steering with a cutoff $x^c \in \mathbb{R}$. Set $s_i = 1$ if $x_i \geq x^c$ and $s_i = 0$ otherwise. Then,

$$\psi(s_i = 1|x_i = x) = \begin{cases} 1 & \text{if } x \geq x^c, \\ 0 & \text{if } x < x^c, \end{cases} \quad \text{and} \quad \psi(s_i = 0|x_i = x) = \begin{cases} 0 & \text{if } x \geq x^c, \\ 1 & \text{if } x < x^c. \end{cases}$$

This signal structure satisfies MLRP, because, for any $x, x' \in \mathbb{R}$ with $x > x'$,

$$\begin{aligned} &\psi(s_i = 1|x_i = x)\psi(s_i = 0|x_i = x') - \psi(s_i = 0|x_i = x)\psi(s_i = 1|x_i = x') \\ &\quad = \begin{cases} 1 & \text{if } x \geq x^c \wedge x' < x^c, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Moreover, if $x \geq x^c > x'$, $\psi(\cdot|x_i = x) \neq \psi(\cdot|x_i = x')$; that is, any binary technology is informative. □

LEMMA 5 (MILGROM, 1981, PROPOSITION 1, SUFFICIENCY). *Suppose that the signal structure satisfies MLRP and is informative. For any $s, s' \in \mathbb{R}$ with $s > s'$, the distribution of $x_i|s_i = s$ (weakly) first-order stochastically dominates the distribution of $x_i|s_i = s'$. By implication, the consumer buys with (weakly) higher probability when being offered a product with signal $s_i = s$ instead of $s_i = s'$.*

PROOF. Consider the case of discretely distributed signals, whose proof is omitted in Milgrom (1981). Because the signal structure satisfies MLRP, for any $x, x' \in \mathbb{R}$ with $x > x'$ and $s, s' \in \mathbb{R}$ with $s > s'$,

$$\psi(s_i = s|x_i = x)\psi(s_i = s'|x_i = x') - \psi(s_i = s'|x_i = x)\psi(s_i = s|x_i = x') \geq 0.$$

This implies that, for any $x^* \geq x'$,

$$\int_{x^*}^{\infty} \psi(s_i = s|x_i = x)\psi(s_i = s'|x_i = x') - \psi(s_i = s'|x_i = x)\psi(s_i = s|x_i = x')\, d\Phi(x) \geq 0.$$

By Bayes' rule, the left-hand side above is equal to

$$\begin{aligned}\psi(s_i = s'|x_i = x') &\int_{x^*}^{\infty} \overbrace{\psi(s_i = s|x_i = x)\phi(x)}^{=\ \psi(s)\phi(x_i=x|s_i=s)}\, dx \\ &- \psi(s_i = s|x_i = x') \int_{x^*}^{\infty} \underbrace{\psi(s_i = s'|x_i = x)\phi(x)}_{=\ \psi(s')\phi(x_i=x|s_i=s')}\, dx.\end{aligned}$$

It follows that

$$\begin{aligned}\psi(s_i = s'|x_i = x')&\psi(s)[1 - \Phi(x_i = x^*|s_i = s)] \\ &- \psi(s_i = s|x_i = x')\psi(s')[1 - \Phi(x_i = x^*|s_i = s')] \\ \geq 0.&\end{aligned}$$

Integrating this, in the same fashion, with respect to $\Phi(x')$ over $(-\infty, x^*]$ and using $\psi(s')$, $\psi(s) > 0$, because both signals are in the support of the unconditional distribution, we obtain

$$[1 - \Phi(x_i = x^*|s_i = s)]\Phi(x_i = x^*|s_i = s') - [1 - \Phi(x_i = x^*|s_i = s')]\Phi(x_i = x^*|s_i = s) \geq 0.$$

This simplifies to

$$\Phi(x_i = x^*|s_i = s') \geq \Phi(x_i = x^*|s_i = s), \tag{A2}$$

which shows that observing a higher signal improves the conditional value distribution in terms of (weak) first-order stochastic dominance. Moreover, because the signal structure is informative, there exists some pair of signals s and s', and some value x^* such that the above inequality is strict.

To see the implication for the purchase probability, consider value-based steering, and take products i and j with signals $s_i > s_j$. (The cases of mistake-based and perceived-value-based steering are analogous.) When being recommended product i, the consumer buys with probability

$$\begin{aligned}\int_{-\infty}^{\infty} 1 - F(\bar{w} - v)\, dG(v|s_i) &= G(v|s_i)[1 - F(\bar{w} - v)]\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f(\bar{w} - v)G(v|s_i)\, dv \\ &= 1 - \int_{-\infty}^{\infty} f(\bar{w} - v)G(v|s_i)\, dv.\end{aligned}$$

Hence, the change in purchase probability when recommending product i instead of j is

$$\int_{-\infty}^{\infty} f(\bar{w} - v) \underbrace{[G(v|s_j) - G(v|s_i)]}_{\geq\ 0 \text{ by (A2)}}\, dv \geq 0.$$

This completes the proof. □

Using the notation introduced in Appendix A.1, we obtain the following lemmas on how (strong) steering changes the distribution of the respective parameter.

LEMMA 6 (POSTERIOR AND PRIOR SATISFY MLRP).

(*i*) *In the case of value-based steering, $g(v|i)$ and $g(v)$ satisfy MLRP and $g(\cdot|i) \neq g(\cdot)$.*
(*ii*) *In the case of mistake-based steering, $f(m|i)$ and $f(m)$ satisfy MLRP and $f(\cdot|i) \neq f(\cdot)$.*
(*iii*) *In the case of perceived-value-based steering, $h(\tilde{v}|i)$ and $h(\tilde{v})$ satisfy MLRP and $h(\cdot|i) \neq h(\cdot)$.*

PROOF. By Lemma 5, it is optimal for the intermediary to recommend a product i with the highest signal. Thus, by assumption, it must do so. Let $s_{-i} := \max\{s_j\}_{j\neq i}$ be the highest signal among other products, and define $m := |\{k : s_k = s_{-i} \text{ for all } k \neq i\}|$ to be the number of other products that have this signal. Let $\Theta(s_{-i}|m)$ be the CDF of s_{-i} conditional on exactly $m \leq I - 1$ products having this signal. The probability that product i is recommended conditional on $x_i = x$ can be written as

$$\mathbb{P}[\text{product } i \text{ is recommended}|x_i = x]$$
$$= \sum_{k=1}^{I-1} \mathbb{P}[m = k] \int \underbrace{1 - \Psi(s|x)}_{= \mathbb{P}[s_i > s_{-i}|x_i = x, s_{-i} = s]} + \frac{1}{k+1} \underbrace{[\Psi(s|x) - \Psi^-(s|x)]}_{= \mathbb{P}[s_i = s_{-i}|x_i = x, s_{-i} = s]} d\Theta(s|m = k).$$

When x_i changes from x' to $x > x'$, the probability of recommending product i changes by

$$\mathbb{P}[\text{product } i \text{ is recommended}|x_i = x] - \mathbb{P}[\text{product } i \text{ is recommended}|x_i = x']$$
$$= \sum_{k=1}^{I-1} \mathbb{P}[m = k] \int \frac{k}{k+1} \underbrace{[\Psi(s|x') - \Psi(s|x)]}_{\geq 0 \text{ by MLRP and Lemma 2}}$$
$$+ \frac{1}{k+1} \underbrace{[\Psi^-(s|x') - \Psi^-(s|x)]}_{\geq 0 \text{ by MLRP and Lemma 2}} d\Theta(s|m = k) \geq 0.$$

Furthermore, because the signal structure is informative, there exists some pair of values x and x' for which the above inequality is strict. Hence, the probability of product i being recommended conditional on $x_i = x$ is weakly increasing in x everywhere and strictly so somewhere.

Consider value-based steering. By Bayes' rule and symmetry across products, respectively,

$$\frac{g(v|i)}{g(v)} = \frac{\mathbb{P}[\text{product } i \text{ is recommended}|v_i = v]}{\mathbb{P}[\text{product } i \text{ is recommended}]} = I\,\mathbb{P}[\text{product } i \text{ is recommended}|v_i = v].$$

By the arguments above, this is weakly increasing in v everywhere and strictly so somewhere. Thus, $g(\cdot|i) \neq g(\cdot)$ and $g(v|i)$ and $g(v)$ satisfy MLRP. By an analogous argument, $f(m|i)/f(m)$ and $h(\tilde{v}|i)/h(\tilde{v})$ are increasing and differ from one somewhere, which proves statements (ii) and (iii). □

LEMMA 7 (STRONG VALUE-BASED AND MISTAKE-BASED STEERING). *Fix any G, F and $\tilde{v}(\cdot)$.*

(*i*) *The probability of purchase under value-based steering approaches 1 if and only if $G(\cdot|i)$ approaches 0 pointwise.*
(*ii*) *The probability of purchase under mistake-based steering approaches 1 if and only if $F(\cdot|i)$ approaches 0 pointwise.*

PROOF. Consider value-based steering; the argument for mistake-based steering is analogous. Take a sequence of CDFs $\{G_n(\cdot|i)\}_{n\in\mathbb{N}}$. If this sequence converges to 0 pointwise then

$$\lim_{n\to\infty}\int_{-\infty}^{\infty} 1 - G_n(\bar{w}-m|i)\,dF(m) = 1 - \int_{-\infty}^{\infty}\lim_{n\to\infty} G_n(\bar{w}-m|i)\,dF(m) = 1,$$

because, by Lebesgue's dominated convergence theorem, we can take the limit inside the integral. Hence, if $G(\cdot|i)$ converges to 0 pointwise, then the probability of purchase approaches 1.

Otherwise, because $\{G_n(\cdot|i)\}_{n\in\mathbb{N}}$ is bounded, by the Bolozano-Wierstraß theorem, there exist a convergent subsequence $\{G_{n'}(\cdot|i)\}_{n'\in\mathbb{N}}$ as well as some $v' \in \mathbb{R}$ and $\epsilon > 0$ such that $\lim_{n'\to\infty} G_{n'}(v'|i) = \epsilon$. Since a CDF is increasing, for any $v \in (v', \infty)$, $\lim_{n'\to\infty} G_{n'}(v|i) \geq \epsilon$. Thus, by Lebesgue's dominated convergence theorem, the probability of purchase cannot converge to 1:

$$\begin{aligned}\lim_{n\to\infty}\int_{-\infty}^{\infty} 1 - G_n(\bar{w}-m|i)\,dF(m) &< 1 - \int_{-\infty}^{\bar{w}-v'}\lim_{n\to\infty} G_n(\bar{w}-m|i)\,dF(m)\\ &\leq\ 1 - \epsilon F(\bar{w}-v')\\ &< 1.\end{aligned}$$

Hence, if $G(\cdot|i)$ does not converge to 0 pointwise, the probability of purchase does not go to 1. □

A.4. Omitted Proofs

PROOF OF PROPOSITION 1. Fix G, F_0 and $\tilde{v}(\cdot)$. Let $m_i \sim F_M$. Because $\tilde{v}(\cdot)$ is fixed, there is some $\bar{w} \in \mathbb{R}$ such that, for any F_M, the consumer buys if and only if $w_i \geq \bar{w}$. By Lebesgue's dominated convergence theorem,

$$\begin{aligned}&\lim_{M\to\infty}\int_{-\infty}^{\infty} v[1 - F_0(\bar{w}-v-M)]\,dG(v)\\ &= \int_{-\infty}^{\infty} v\,dG(v) - \int_{-\infty}^{\infty} v\lim_{M\to\infty} F_0(\bar{w}-v-M)\,dG(v)\\ &= \mathbb{E}[v_i].\end{aligned}$$

Because $\mathbb{E}[v_i] < 0$, the consumer does not buy reasonably for sufficiently large M. And because any consumer satisfies at least one notion of reasonability (see footnote 8), she refrains reasonably. □

PROOF OF REMARK 1. The proof follows from steering improving the value distribution in terms of FOSD. □

PROOF OF PROPOSITION 2. *Equilibrium.* Let $S \subseteq \mathbb{R}$ and $W = \mathbb{R}$ be the signal spaces of the intermediary and consumer, respectively. A strategy of the intermediary is a mapping $\sigma_F : S^I \to \Delta(\{1, 2, \ldots, I\})$, with the interpretation that each product is recommended with the corresponding probability. The consumer's strategy $\sigma_C : \mathbb{R} \to [0, 1]$ is a mapping from her signal regarding the recommended product to her purchase probability. A pair of strategies (σ_F, σ_C) is a BNE if each player best responds to the other player's strategy for each realisation of her signal.

(i) Consider value-based steering. First, we argue that *if the intermediary recommends the product with the highest value, the consumer best responds by using a(n improper) cutoff strategy* *in which she buys if her signal exceeds some cutoff* $\tilde{w} \in \overline{\mathbb{R}}$ *and refrains if it falls below* $\tilde{w}$. Because under value-based steering the probability of product i being recommended is independent of m_i,

$$\begin{aligned}\mathbb{P}[w_i \le w | i \text{ recommended}, v_i = v] &= \mathbb{P}[m_i \le w - v | i \text{ recommended}] \\ &= \mathbb{P}[m_i \le w - v] \\ &= F(w - v).\end{aligned}$$

Hence, the likelihood ratio of observing $w_i = w$ given $v_i = v$ and $v_i = v' < v$, respectively, is $f(w - v)/f(w - v')$. Substituting $m = w - v$ and $x = v - v' > 0$, this becomes $f(m)/f(m + x)$, which increases in m, and thus w, by footnote 2. Hence, if the intermediary recommends the product with the highest signal, the consumer's (endogenously arising) signal structure satisfies MLRP; thus, $\mathbb{E}[v_i | i \text{ recommended}, w_i = w]$ is increasing in w. Hence, the consumer uses a cutoff strategy.

Second, *for any given consumer cutoff strategy, it is optimal for the intermediary to recommend a product for which it observed the highest signal.* Consider products i and j with signals $s_i > s_j$. By Lemma 5, the consumer is weakly more likely to buy product i. It is thus optimal for the intermediary to recommend a product with the highest signal. Hence, such a BNE exists.

Third, we argue that *a rational consumer benefits from value-based steering.* Absent steering, a rational consumer behaves as an always reasonable consumer does. With steering, a rational consumer differs from an always reasonable consumer only in that she may use a different cutoff. If she adopted the same strategy as her strategically naive counterpart, by Proposition 3, her welfare would increase through steering. Adopting the optimal strategy instead (weakly) increases her welfare, so a rational consumer benefits from value-based steering.

(ii) Consider perceived-value-based steering. Using Bayes' rule and the chain rule,

$$\begin{aligned}\mathbb{P}[v_i \ge v | i \text{ recommended}, w_i = w] &= \mathbb{P}[v_i \ge v | w_i = w] \frac{\mathbb{P}[i \text{ recommended} | v_i \ge v, w_i = w]}{\mathbb{P}[i \text{ recommended} | w_i = w]} \\ &= \mathbb{P}[v_i \ge v | w_i = w],\end{aligned}$$

where the second equality follows from the fact that, under perceived-value-based steering, when conditioning on w_i, the recommendation is independent of v_i. Thus, the consumer uses the same cutoff strategy with and without steering. More specifically, she behaves identical to an always reasonable consumer who is strategically naive.

Taking the consumer's cutoff strategy as given, by Lemma 5, the intermediary cannot do better than recommending a product for which it observed the highest signal. Hence, such a BNE exists. And because the rational consumer behaves identical to her strategically naive counterpart, by Proposition 5, she benefits from perceived-value-based steering.

(iii) Consider perfect mistake-based steering. We first show that *if the intermediary recommends the product with the highest mistake, the consumer best responds by using a(n improper) cutoff strategy.* The consumer's mistake for the suggested product i^* is distributed according to the CDF

$$\mathbb{P}[\max\{m_1, \ldots, m_I\} \le m] = \mathbb{P}[m_1 \le m]^I = F(m)^I$$

with density $If(m)F(m)^{I-1}$. By the same argument as in footnote 2, the rational consumer's (endogenously arising) signal structure satisfies MLRP if and only if, for any $x > 0$,

$$\frac{If(m)F(m)^{I-1}}{If(m + x)F(m + x)^{I-1}} = \frac{f(m)}{f(m + x)} \left(\frac{F(m)}{F(m + x)}\right)^{I-1}$$

is increasing in m. Because absent steering the consumer's signal structure satisfies MLRP, $f(m)/f(m+x)$ is increasing in m. Moreover, $F(m)/F(m+x)$ is increasing in m if and only if

$$f(m)F(m+x) - F(m)f(m+x) \geq 0 \quad \text{or, equivalently,} \quad \frac{f(m)}{F(m)} \geq \frac{f(m+x)}{F(m+x)};$$

that is, if and only if F is log-concave. This follows from our assumption that f is log-concave (e.g., Bagnoli and Bergstrom, 2005). Thus, if the intermediary recommends the product with the highest mistake, the consumer's signal structure satisfies MLRP, so $\mathbb{E}[v_i|m_i = \max\{m_1, \ldots, m_I\}, w_i = w]$ is increasing in w. Hence, the consumer uses a cutoff strategy.

Second, *if the consumer uses a cutoff strategy with a cutoff $\tilde{w} \in \overline{\mathbb{R}}$, the intermediary best responds by recommending the product with the highest mistake.* Consider products i and j with mistakes $m_i = m > m_j$. When recommending i, the consumer buys product with probability $1 - G(\tilde{w} - m)$, which is weakly greater than the purchase probability when recommending j. Hence, a BNE in which the intermediary recommends the product with the highest mistake exists.

Third, we show that *the consumer benefits from perfect mistake-based steering for sufficiently large I*. To do so, we study the distribution of $m^s := \max\{m_1, \ldots, m_I\}$. Define $\bar{m}_I$ implicitly by

$$F(\bar{m}_I) = 1 - \frac{1}{I}, \tag{A3}$$

which is well defined because m_i has full support, and which implies that $\lim_{I\to\infty} \bar{m}_I = \infty$. Because mistakes are independently distributed across products with a hazard rate that approaches infinity, by Hansen (2020), there exists some $\hat{m} \in \mathbb{R}$ such that, for any $m' > \hat{m}$, one has

$$\mathbb{P}[m^s \leq m'] = \left[1 - \frac{e^{-r(m')}}{I}\right]^I \quad \text{with} \quad r(m') := If(\bar{m}_I)(m' - \bar{m}_I). \tag{A4}$$

Because $\mathbb{P}[m^s \leq \bar{m}_I - v] \in [0, 1]$ for any $v \leq \bar{m}_I + \hat{m}$ (so that (A4) applies), $e^{vIf(\bar{m}_I)}/I \in [0, 1]$ also. Furthermore, using l'Hôpital's rule and $d\bar{m}_I/dI = 1/(I^2 f(\bar{m}_I))$, we have, for all $v \leq \bar{m}_I + \hat{m}$,

$$\lim_{I\to\infty} \frac{e^{vIf(\bar{m}_I)}}{I} = \lim_{I\to\infty} v \underbrace{\left[If(\bar{m}_I) + \frac{f'(\bar{m}_I)}{f(\bar{m}_I)}\right]}_{(\star)} \frac{e^{vIf(\bar{m}_I)}}{I},$$

which we now show requires $\lim_{I\to\infty} e^{vIf(\bar{m}_I)}/I = 0$. If the limit of $(\star)$ is zero, this follows from $e^{vIf(\bar{m}_I)}/I \in [0, 1]$. Otherwise, the equality can also only hold for *all* such v if $\lim_{I\to\infty} e^{vIf(\bar{m}_I)}/I = 0$.

Next, re-arranging (A3) gives

$$If(\bar{m}_I) = \frac{f(\bar{m}_I)}{1 - F(\bar{m}_I)}. \tag{A5}$$

Hence, because $\lim_{I\to\infty} \bar{m}_I = \infty$ and because the hazard rate of m_i diverges, $\lim_{I\to\infty} If(\bar{m}_I) = \infty$. Using (A5) and $d\bar{m}_I/dI = 1/(I^2 f(\bar{m}_I))$, we have

$$If(\bar{m}_I) + \frac{f'(\bar{m}_I)}{f(\bar{m}_I)} = I \frac{\partial}{\partial I} If(\bar{m}_I) = I \frac{\partial}{\partial I} \frac{f(\bar{m}_I)}{1 - F(\bar{m}_I)} = I \underbrace{\frac{d\bar{m}_I}{dI}}_{>0} \underbrace{\frac{\partial}{\partial m} \frac{f(m)}{1 - F(m)}\bigg|_{m=\bar{m}_I}}_{>0} > 0.$$

Using l'Hôpital's rule and $d\bar{m}_I/dI = 1/(I^2 f(\bar{m}_I))$,

$$\begin{aligned}\lim_{I\to\infty} \ln\left(\left[1 - \frac{e^{vIf(\bar{m}_I)}}{I}\right]^I\right) &= \lim_{I\to\infty} \frac{\ln[1 - e^{vIf(\bar{m}_I)}/I]}{1/I} \\ &= \lim_{I\to\infty} \frac{e^{vIf(\bar{m}_I)}[v(If(\bar{m}_I) + f'(\bar{m}_I)/f(\bar{m}_I)) - 1]}{1 - e^{vIf(\bar{m}_I)}/I} \\ &= -\lim_{I\to\infty} e^{vIf(\bar{m}_I)}, \qquad \text{(A6)}\end{aligned}$$

where the last equality uses $\lim_{I\to\infty} e^{vIf(\bar{m}_I)}/I = 0$ and the fact that (as we show next) $If(\bar{m}_I) + f'(\bar{m}_I)/f(\bar{m}_I)$ goes to zero as $I \to \infty$. Suppose, for the sake of a contradiction, that $If(\bar{m}_I) + f'(\bar{m}_I)/f(\bar{m}_I)$ does not go to zero, and denote the limit by $l > 0$. Then, for any $v > 1/l$, $\lim_{I\to\infty} v(If(\bar{m}_I) + f'(\bar{m}_I)/f(\bar{m}_I)) > 1$. Because $\lim_{I\to\infty} If(\bar{m}_I) = \infty$, this implies that, for any $v > 1/l$, $\lim_{I\to\infty} \ln([1 - e^{vIf(\bar{m}_I)}/I]^I) = \infty$; a contradiction. We conclude that $If(\bar{m}_I) + f'(\bar{m}_I)/f(\bar{m}_I)$ goes to zero as $I \to \infty$.

Combining $\lim_{I\to\infty} If(\bar{m}_I) = \infty$ with (A6), we have, for all $v > 0$,

$$\lim_{I\to\infty} \ln\left(\left[1 - \frac{e^{vIf(\bar{m}_I)}}{I}\right]^I\right) = -\infty \quad \text{or, equivalently,} \quad \lim_{I\to\infty}\left[1 - \frac{e^{vIf(\bar{m}_I)}}{I}\right]^I = 0, \qquad \text{(A7)}$$

while, for all $v < 0$,

$$\lim_{I\to\infty} \ln\left(\left[1 - \frac{e^{vIf(\bar{m}_I)}}{I}\right]^I\right) = 0 \quad \text{or, equivalently,} \quad \lim_{I\to\infty}\left[1 - \frac{e^{vIf(\bar{m}_I)}}{I}\right]^I = 1. \qquad \text{(A8)}$$

We now provide a consumer cutoff strategy that, as $I \to \infty$, brings her arbitrarily close to her first-best welfare. Let I be large enough so that $\bar{m}_I + \hat{m} > 0$. Consider a consumer who buys the recommended product i^* if and only if $w_{i^*} > \bar{m}_I$. Using (A4), her welfare is

$$\begin{aligned}&\int_{-\infty}^{\infty} v\mathbb{P}[v + m^s > \bar{m}_I]\, dG(v) \\ &\quad = \int_{\bar{m}_I+\hat{m}}^{\infty} v\mathbb{P}[m^s > \bar{m}_I - v]\, dG(v) \\ &\qquad + \int_0^{\bar{m}_I+\hat{m}} v\left[1 - \left[1 - \frac{e^{vIf(\bar{m}_I)}}{I}\right]^I\right] dG(v) + \int_{-\infty}^{0} v\left[1 - \left[1 - \frac{e^{vIf(\bar{m}_I)}}{I}\right]^I\right] dG(v) \\ &\quad \geq \int_0^{\bar{m}_I+\hat{m}} v\left[1 - \left[1 - \frac{e^{vIf(\bar{m}_I)}}{I}\right]^I\right] dG(v) + \int_{-\infty}^{0} v\left[1 - \left[1 - \frac{e^{vIf(\bar{m}_I)}}{I}\right]^I\right] dG(v).\end{aligned}$$

Using (A7), (A8) and $\lim_{I\to\infty} \bar{m}_I = \infty$, in the limit as $I \to \infty$ the lower bound above approaches $\int_0^{\infty} v\, dG(v)$. This is the first-best consumer welfare and, hence, for sufficiently large I, perfect mistake-based steering benefits a rational consumer. □

PROOF OF PROPOSITION 3. By Lemma 6, $g(\cdot|i)$ and $g(\cdot)$ satisfy MLRP and $g(\cdot|i) \neq g(\cdot)$. Thus, by Lemma 1, there exists some $v^* \in \mathbb{R}$ such that $g(v|i)/g(v) > 1$ if $v > v^*$ and $g(v|i)/g(v) < 1$ if $v < v^*$.

(i) We first show that, *as the purchase probability goes to* 1, $v^* > 0$. Suppose otherwise. For any $V > 0$, as the purchase probability goes to 1, by Lemma 7, eventually $G(V|i) < G(V) - G(0)$.

Thus,

$$\begin{aligned}
G(V) - G(0) &> G(V|i) \\
&> \int_0^V g(v|i)\,dv \\
&= \int_0^V g(v)\,\underbrace{\frac{g(v|i)}{g(v)}}_{\geq 1 \text{ due to } v^* < 0}\,dv \\
&\geq \int_0^V g(v)\,dv \\
&= G(V) - G(0);
\end{aligned}$$

a contradiction. Hence, for sufficiently high purchase probabilities, $v^* > 0$. If $v^* \geq 0$, using (1) and (A1), the change in welfare due to steering is

$$\begin{aligned}
&\int_{-\infty}^{0} \overbrace{v}^{<0}\,\overbrace{[1 - F(\bar{w} - v)]}^{>0}\,\overbrace{[g(v|i) - g(v)]}^{\leq 0 \text{ since } v^* \geq 0}\,dv + \int_0^{\infty} v[1 - F(\bar{w} - v)][g(v|i) - g(v)]\,dv \\
&\quad \geq \int_0^{v^*} v[1 - F(\bar{w} - v)][g(v|i) - g(v)]\,dv + \int_{v^*}^{\infty} v[1 - F(\bar{w} - v)][g(v|i) - g(v)]\,dv \\
&\quad > v^*[1 - F(\bar{w} - v^*)]\left(\int_0^{v^*} [g(v|i) - g(v)]\,dv + \int_{v^*}^{\infty} g(v|i) - g(v)\,dv\right) \\
&\quad = \underbrace{v^*[1 - F(\bar{w} - v^*)]}_{\geq 0 \text{ by } v^* \geq 0}\,\underbrace{[G(0) - G(0|i)]}_{>0 \text{ by Lemma 2 (b)}} \\
&\quad \geq 0.
\end{aligned}$$

Hence, steering that induces the consumer to buy with a sufficiently high probability benefits her. Therefore, by definition, sufficiently strong steering benefits the consumer.

(ii) Suppose that the consumer does not buy reasonably. Consider binary steering with a cutoff $v^c \in \mathbb{R}$, and let $v_{\max}$ be the largest v_i among the I products. Welfare under steering is then

$$\begin{aligned}
&\mathbb{P}[v_{\max} \geq v^c] \int_{v^c}^{\infty} v[1 - F(\bar{w} - v)]\,dG(v|v \geq v^c) \\
&\qquad + \mathbb{P}[v_{\max} < v^c] \int_{-\infty}^{v^c} v[1 - F(\bar{w} - v)]\,dG(v|v < v^c) \\
&\quad = \frac{1 - G(v^c)^I}{1 - G(v^c)} \int_{v^c}^{\infty} v[1 - F(\bar{w} - v)]\,dG(v) + \frac{G(v^c)^I}{G(v^c)} \int_{-\infty}^{v^c} v[1 - F(\bar{w} - v)]\,dG(v).
\end{aligned}$$

Subtracting the welfare absent steering, $\int_{-\infty}^{\infty} v[1 - F(\bar{w} - v)]\,dG(v)$, from the above yields

$$\frac{1 - G(v^c)^{I-1}}{1 - G(v^c)}\left[G(v^c)\int_{-\infty}^{\infty} v[1 - F(\bar{w} - v)]\,dG(v) - \int_{-\infty}^{v^c} v[1 - F(\bar{w} - v)]\,dG(v)\right]. \tag{A9}$$

Since I does not affect the sign of the welfare effect, we set $I = 2$, so that (A9) simplifies to

$$\Delta(v^c) := G(v^c)\int_{-\infty}^{\infty} v[1 - F(\bar{w} - v)]\,dG(v) - \int_{-\infty}^{v^c} v[1 - F(\bar{w} - v)]\,dG(v).$$

Taking the first derivative of the above expression (with respect to v^c) yields

$$\Delta'(v^c) = g(v^c)\left[\int_{-\infty}^{\infty} v[1 - F(\bar{w} - v)]\,dG(v) - v^c[1 - F(\bar{w} - v^c)]\right]. \tag{A10}$$

Since the signal becomes uninformative, $\Delta(v^c) \to 0$ as $v^c \to -\infty$. Thus, it is sufficient to show that $\Delta'(v^c) < 0$ for sufficiently small v^c. Since the consumer does not buy reasonably, by (A10), it suffices to verify that $\lim_{v^c \to -\infty} v^c[1 - F(\bar{w} - v^c)] = 0$ or, equivalently, $\lim_{v' \to \infty}(\bar{w} + v')[1 - F(\bar{w} + v')] = 0$. Let $v' > -\bar{w}$. Since the expectation of m_i exists and is finite, $\lim_{v' \to \infty} \int_0^{\bar{w}+v'} mf(m)\,dm < \infty$.[28] And since this integral is increasing in v', $\lim_{v' \to \infty} \int_{\bar{w}+v'}^{\infty} mf(m)\,dm = 0$. At the same time,

$$\int_{\bar{w}+v'}^{\infty} mf(m)\,dm > (\bar{w} + v')\int_{\bar{w}+v'}^{\infty} f(m)\,dm = (\bar{w} + v')[1 - F(\bar{w} + v')] \geq 0,$$

so that $\lim_{v' \to \infty}(\bar{w} + v')[1 - F(\bar{w} + v')] = 0$, which was to be proven.

(iii) '(a) $\Rightarrow$ (b)'. Suppose that the consumer buys reasonably. By the proof of part (i), any value-based steering raises welfare whenever $v^* \geq 0$. For any $v^* < 0$, because steering improves the value distribution in terms of MLRP, the change in welfare is given by

$$\begin{aligned}
&\int_{-\infty}^{v^*} \overbrace{v}^{<0}\, \overbrace{[1 - F(\bar{w} - v)]}^{>0}\, \overbrace{[g(v|i) - g(v)]}^{\leq 0}\, dv + \int_{v^*}^{\infty} v[1 - F(\bar{w} - v)][g(v|i) - g(v)]\, dv \\
&\quad \geq \int_{\min\{v^*,0\}}^{\infty} v[1 - F(\bar{w} - v)][g(v|i) - g(v)]\, dv \\
&\quad = \int_{\min\{v^*,0\}}^{0} v[1 - F(\bar{w} - v)]\left[\frac{g(v|i)}{g(v)} - 1\right] dG(v) \\
&\qquad + \int_{0}^{\infty} v[1 - F(\bar{w} - v)]\left[\frac{g(v|i)}{g(v)} - 1\right] dG(v) \\
&\quad \geq \int_{\min\{v^*,0\}}^{0} v[1 - F(\bar{w} - v)]\left[\frac{g(0|i)}{g(0)} - 1\right] dG(v) \\
&\qquad + \int_{0}^{\infty} v[1 - F(\bar{w} - v)]\left[\frac{g(0|i)}{g(0)} - 1\right] dG(v) \\
&\quad = \left[\frac{g(0|i)}{g(0)} - 1\right]\int_{\min\{v^*,0\}}^{\infty} v[1 - F(\bar{w} - v)]\, dG(v) \\
&\quad > \left[\frac{g(0|i)}{g(0)} - 1\right]\int_{-\infty}^{\infty} v[1 - F(\bar{w} - v)]\, dG(v) \\
&\quad > 0,
\end{aligned}$$

[28] Because the distribution of m_i has an increasing hazard rate, it also has finite moments (e.g., Barlow *et al.*, 1963, p. 382). And, if a random variable X has a finite expectation, both its positive part, $\max\{0, X\}$, and its negative part, $-\min\{0, X\}$, are integrable (e.g., https://en.wikipedia.org/wiki/Expected_value, 'Basic properties').

where the last inequality follows from the consumer buying reasonably and $v^* < 0$.

'(b) $\Rightarrow$ (a)': follows from part (ii).

(iv) The proof follows from part (iii) because both types of consumers buy reasonably. □

PROOF OF PROPOSITION 4. By Lemma 6, $f(\cdot|i)$ and $f(\cdot)$ satisfy MLRP and $f(\cdot|i) \neq f(\cdot)$. Thus, by Lemma 1, there exists $m^* \in \mathbb{R}$ such that $f(m|i)/f(m) > 1$ if $m > m^*$ and $f(m|i)/f(m) < 1$ if $m > m^*$.

(ii) Suppose that the consumer buys reasonably. Consider binary steering with a cutoff $m^c \in \mathbb{R}$, and let $m_{\max}$ be the largest m_i among the I products. In this case, welfare under steering is

$$\begin{aligned}
&\mathbb{P}[m_{\max} \geq m^c] \int_{-\infty}^{\infty} v \int_{m^c}^{\infty} \mathbb{1}_{\{v+m\geq\bar{w}\}}\, dF(m|m \geq m^c)\, dG(v) \\
&\quad + \mathbb{P}[m_{\max} < m^c] \int_{-\infty}^{\infty} v \int_{-\infty}^{m^c} \mathbb{1}_{\{v+m\geq\bar{w}\}}\, dF(m|m < m^c)\, dG(v) \\
&= [1 - F(m^c)^I]\left[\int_{-\infty}^{\bar{w}-m^c} v \frac{1-F(\bar{w}-v)}{1-F(m^c)}\, dG(v) + \int_{\bar{w}-m^c}^{\infty} v \frac{1-F(m^c)}{1-F(m^c)}\, dG(v)\right] \\
&\quad + F(m^c)^I \int_{\bar{w}-v}^{\infty} v \frac{F(m^c) - F(\bar{w}-v)}{F(m^c)}\, dG(v) \\
&= \frac{1-F(m^c)^I}{1-F(m^c)} \int_{-\infty}^{\bar{w}-m^c} v[1-F(\bar{w}-v)]\, dG(v) \\
&\quad + \int_{\bar{w}-m^c}^{\infty} v\, dG(v) - F(m^c)^{I-1} \int_{\bar{w}-m^c}^{\infty} v F(\bar{w}-v)\, dG(v).
\end{aligned}$$

The welfare absent steering is given by

$$\begin{aligned}
\int_{-\infty}^{\infty} v[1-F(\bar{w}-v)]\, dF(v) &= \int_{-\infty}^{\bar{w}-m^c} v[1-F(\bar{w}-v)]\, dG(v) \\
&\quad + \int_{\bar{w}-m^c}^{\infty} v\, dG(v) - \int_{\bar{w}-m^c}^{\infty} v F(\bar{w}-v)\, dG(v).
\end{aligned}$$

The welfare effect of mistake-based steering is thus

$$[1 - F(m^c)^{I-1}]\left[\frac{F(m^c)}{1-F(m^c)} \int_{-\infty}^{\bar{w}-m^c} v[1-F(\bar{w}-v)]\, dG(v) + \int_{\bar{w}-m^c}^{\infty} v F(\bar{w}-v)\, dG(v)\right].$$

The sign of the welfare effect is independent of the number of products. Hence, it suffices to consider $I = 2$. In this case, the welfare effect of marginally increasing m^c is

$$f(m^c)\left[\int_{-\infty}^{\bar{w}-m^c} v[1-F(\bar{w}-v)]\, dG(v) - \int_{\bar{w}-m^c}^{\infty} v F(\bar{w}-v)\, dG(v)\right].$$

Since the signal becomes uninformative, the change in welfare approaches zero as $m^c \to -\infty$. It thus suffices to verify that, for sufficiently small m^c, the above derivative is strictly positive:

$$\begin{aligned}&\lim_{m^c\to-\infty}\left\{\int_{-\infty}^{\bar{w}-m^c} v[1-F(\bar{w}-v)]\,dG(v)-\int_{\bar{w}-m^c}^{\infty} vF(\bar{w}-v)\,dG(v)\right\}\\&=\int_{-\infty}^{\infty} v[1-F(\bar{w}-v)]\,dG(v)\\&>0.\end{aligned}$$

(i) '(a) $\Rightarrow$ (b)'. The welfare effect of mistake-based steering can be written as

$$\begin{aligned}&\int_{-\infty}^{\infty} v[F(\bar{w}-v)-F(\bar{w}-v|i)]\,dG(v)\\&=\int_{-\infty}^{0} \overbrace{v[1-F(\bar{w}-v)]}^{<0}\,\overbrace{\frac{F(\bar{w}-v)-F(\bar{w}-v|i)}{1-F(\bar{w}-v)}}^{\geq 0}\,dG(v)\\&\quad+\int_{0}^{\infty} \underbrace{v[1-F(\bar{w}-v)]}_{>0}\,\underbrace{\frac{F(\bar{w}-v)-F(\bar{w}-v|i)}{1-F(\bar{w}-v)}}_{\geq 0}\,dG(v).\end{aligned}$$

By Lemmas 2(b) and 3(a), for any $v\in(-m^*+\bar{w},\infty)$, not only is $[F(\bar{w}-v)-F(\bar{w}-v|i)]/[1-F(\bar{w}-v)]$ *strictly* positive, but also *strictly* decreasing in v. Hence, the above is *strictly* smaller than

$$\begin{aligned}&\int_{-\infty}^{0} v[1-F(\bar{w}-v)]\frac{F(\bar{w})-F(\bar{w}|i)}{1-F(\bar{w})}\,dG(v)+\int_{0}^{\infty} v[1-F(\bar{w}-v)]\frac{F(\bar{w})-F(\bar{w}|i)}{1-F(\bar{w})}\,dG(v)\\&=\frac{F(\bar{w})-F(\bar{w}|i)}{1-F(\bar{w})}\int_{-\infty}^{\infty} v[1-F(\bar{w}-v)]\,dG(v).\end{aligned}$$

Hence, if the consumer does not buy reasonably, mistake-based steering strictly lowers her welfare.

'(b) $\Rightarrow$ (a)': follows by part (ii).

(iv) We prove the statement for a consumer who refrains reasonably; the statement for an always reasonable consumer then follows. Fix $M>\bar{w}$, and let

$$\epsilon := \frac{1}{n}\frac{\int_{-\infty}^{\infty} vF(\bar{w}-v)\,dG(v)}{\int_{\bar{w}-M}^{0} v\,dG(v)}>0,$$

for $n\in\mathbb{N}_{\geq 2}$ large enough such that $\epsilon<1$. Consider steering that induces the consumer to buy with a probability of at least $\underline{\pi}(M):=1-\epsilon G(\bar{w}-M)$, which is increasing in M. Then,

$$1-\epsilon G(\bar{w}-M)<\int_{-\infty}^{\infty} 1-F(\bar{w}-v|i)\,dG(v)<1-F(M|i)G(\bar{w}-M),$$

and thus $F(M|i) < \epsilon$. Hence, fixing any $M > \bar{w}$, the change in welfare satisfies

$$
\begin{aligned}
&\int_{-\infty}^{\infty} v[F(\bar{w}-v) - F(\bar{w}-v|i)]\,dG(v) \\
&\quad < \int_{-\infty}^{\infty} vF(\bar{w}-v)\,dG(v) - \int_{\bar{w}-M}^{0} vF(\bar{w}-v|i)\,dG(v) \\
&\qquad - \int_{-\infty}^{\bar{w}-M} vF(\bar{w}-v|i)\,dG(v) \\
&\quad < \int_{-\infty}^{\infty} vF(\bar{w}-v)\,dG(v) - \epsilon\int_{\bar{w}-M}^{0} v\,dG(v) - \int_{-\infty}^{\bar{w}-M} vF(\bar{w}-v|i)\,dG(v) \\
&\quad = \frac{n-1}{n}\int_{-\infty}^{\infty} vF(\bar{w}-v)\,dG(v) - \int_{-\infty}^{\bar{w}-M} vF(\bar{w}-v|i)\,dG(v),
\end{aligned}
$$

where the second inequality holds since $F(m|i) < \epsilon$ for any $m \leq M$. The last expression above is continuous and strictly decreasing in M. Moreover, as $M \to \infty$, it converges to

$$
\frac{n-1}{n}\int_{-\infty}^{\infty} vF(\bar{w}-v)\,dG(v),
$$

which is strictly negative because the consumer refrains reasonably. Hence, there exists some $\bar{M} \in \mathbb{R}$ such that any steering that induces a purchase probability of at least $\underline{\pi}(\bar{M})$ harms the consumer.

(iii) '(a) $\Rightarrow$ (b)'. Let the consumer not refrain reasonably. The welfare effect of steering is

$$
\begin{aligned}
&\int_{-\infty}^{\infty} v[F(\bar{w}-v) - F(\bar{w}-v|i)]\,dG(v) \\
&= \int_{-\infty}^{0} \overbrace{vF(\bar{w}-v)}^{<0}\,\frac{\overbrace{F(\bar{w}-v) - F(\bar{w}-v|i)}^{\geq 0}}{F(\bar{w}-v)}\,dG(v) \\
&\quad + \int_{0}^{\infty} \underbrace{vF(\bar{w}-v)}_{>0}\,\underbrace{\frac{F(\bar{w}-v) - F(\bar{w}-v|i)}{F(\bar{w}-v)}}_{\geq 0}\,dG(v) \\
&> \int_{-\infty}^{0} vF(\bar{w}-v)\frac{F(\bar{w}) - F(\bar{w}|i)}{F(\bar{w})}\,dG(v) + \int_{0}^{\infty} vF(\bar{w}-v)\frac{F(\bar{w}) - F(\bar{w}|i)}{F(\bar{w})}\,dG(v) \\
&= \frac{F(\bar{w}) - F(\bar{w}|i)}{F(\bar{w})}\int_{-\infty}^{\infty} vF(\bar{w}-v)\,dG(v) \\
&\geq 0,
\end{aligned}
$$

where the first inequality holds weakly by Lemmas 2(b) and 3(b), and the second inequality follows since the consumer does not refrain reasonably. Furthermore, the first inequality is strict because, by Lemmas 2(b) and 3(b), there exists some $\epsilon > 0$ such that, for any $v \in (-(m^* + \epsilon) + \bar{w}, -m^* + \bar{w})$, not only is $[F(\bar{w}-v) - F(\bar{w}-v|i)]/F(\bar{w}-v)$ *strictly* positive, but also *strictly* increasing in v.

'(b) $\Rightarrow$ (a)': follows from part (iv). □

PROOF OF PROPOSITION 5. (i) Using the notation introduced in Appendix A.1, the welfare effect of perceived-value based steering is given by

$$\int_0^\infty [h(\tilde{v}|i) - h(\tilde{v})]\mathbb{E}[v_i|\tilde{v}_i = \tilde{v}]\, d\tilde{v} = \int_0^\infty \left[\frac{h(\tilde{v}|i)}{h(\tilde{v})} - 1\right]\mathbb{E}[v_i|\tilde{v}_i = \tilde{v}]\, dH(\tilde{v}).$$

Since the consumer's signal structure satisfies MLRP and $\tilde{v}(\cdot)$ is strictly increasing, by Proposition 1 of Milgrom (1981), there exists some $\tilde{v}' \in \mathbb{R} \cup \{-\infty, \infty\}$ such that $\mathbb{E}[v_i|\tilde{v}_i = \tilde{v}] > 0$ if and only if $\tilde{v} > \tilde{v}'$. By Lemmas 1 and 6, there exists some $\tilde{v}'' \in \mathbb{R}$ such that $h(\tilde{v}|i)/h(\tilde{v}) > 1$ if and only if $\tilde{v} > \tilde{v}''$.

'(a) $\Rightarrow$ (b)'. Suppose that the consumer buys reasonably. This implies that $\mathbb{E}[v_i|\tilde{v}_i = \tilde{v}] > 0$ eventually and, thus, $\tilde{v}' < \infty$. If $\tilde{v}'' < \tilde{v}'$, we re-write the change in welfare due to steering as

$$\begin{aligned}
&\int_0^{\max\{0,\tilde{v}''\}} \overbrace{\left[\frac{h(\tilde{v}|i)}{h(\tilde{v})} - 1\right]}^{\leq 0} \overbrace{\mathbb{E}[v_i|\tilde{v}_i = \tilde{v}]}^{<0}\, dH(\tilde{v})\\
&\quad + \int_{\max\{0,\tilde{v}''\}}^{\max\{0,\tilde{v}'\}} \overbrace{\left[\frac{h(\tilde{v}|i)}{h(\tilde{v})} - 1\right]}^{>0} \overbrace{\mathbb{E}[v_i|\tilde{v}_i = \tilde{v}]}^{\leq 0}\, dH(\tilde{v})\\
&\quad + \int_{\max\{0,\tilde{v}'\}}^{\infty} \underbrace{\left[\frac{h(\tilde{v}|i)}{h(\tilde{v})} - 1\right]}_{>0} \underbrace{\mathbb{E}[v_i|\tilde{v}_i = \tilde{v}]}_{>0}\, dH(\tilde{v}).
\end{aligned}$$

Because the first integral is (weakly) positive and because $h(\tilde{v}|i)/h(\tilde{v})$ is (weakly) increasing in $\tilde{v}$ due to MLRP, the expression above is bounded from below by

$$\begin{aligned}
&\int_{\max\{0,\tilde{v}''\}}^{\max\{0,\tilde{v}'\}} \left[\frac{h(\max\{0,\tilde{v}'\}|i)}{h(\max\{0,\tilde{v}'\})} - 1\right]\mathbb{E}[v_i|\tilde{v}_i = \tilde{v}]\, dH(\tilde{v})\\
&\quad + \int_{\max\{0,\tilde{v}'\}}^{\infty} \left[\frac{h(\max\{0,\tilde{v}'\}|i)}{h(\max\{0,\tilde{v}'\})} - 1\right]\mathbb{E}[v_i|\tilde{v}_i = \tilde{v}]\, dH(\tilde{v})\\
&= \left[\frac{h(\max\{0,\tilde{v}'\}|i)}{h(\max\{0,\tilde{v}'\})} - 1\right]\int_{\max\{0,\tilde{v}''\}}^{\infty} \mathbb{E}[v_i|\tilde{v}_i = \tilde{v}]\, dH(\tilde{v})\\
&\geq \underbrace{\left[\frac{h(\max\{0,\tilde{v}'\}|i)}{h(\max\{0,\tilde{v}'\})} - 1\right]}_{>0}\underbrace{\int_0^{\infty} \mathbb{E}[v_i|\tilde{v}_i = \tilde{v}]\, dH(\tilde{v})}_{>0}\\
&> 0,
\end{aligned}$$

where the weak inequality follows since $\mathbb{E}[v_i|\tilde{v}_i = \tilde{v}] \leq 0$ for any $\tilde{v} < \tilde{v}'$, and thus for any $\tilde{v} \leq \tilde{v}''$, and the strict inequality holds because the consumer buys reasonably and $\max\{0, \tilde{v}'\} > \tilde{v}''$.

Now let $\tilde{v}'' \geq \tilde{v}'$. If $\tilde{v}'' \leq 0$ then, for any $\epsilon > 0$, the change in welfare due to steering is

$$\int_0^\infty [h(\tilde{v}|i) - h(\tilde{v})]\mathbb{E}[v_i|\tilde{v}_i = \tilde{v}]\, d\tilde{v} \geq \int_\epsilon^\infty \underbrace{[h(\tilde{v}|i) - h(\tilde{v})]}_{>0 \text{ since } \tilde{v}''\leq 0} \underbrace{\mathbb{E}[v_i|\tilde{v}_i = \tilde{v}]}_{>0 \text{ since } \tilde{v}'\leq 0}\, d\tilde{v} > 0.$$

Now let $\tilde{v}'' > 0$. For $M > \mathbb{E}[v_i|\tilde{v}_i = \tilde{v}'']$, the change in welfare due to steering is greater than

$$\int_0^{\max\{0,\tilde{v}'\}} \overbrace{[h(\tilde{v}|i) - h(\tilde{v})]}^{<0} \overbrace{\mathbb{E}[v_i|\tilde{v}_i = \tilde{v}]}^{\leq 0} \, d\tilde{v}$$

$$+ \int_{\max\{0,\tilde{v}'\}}^{\infty} [h(\tilde{v}|i) - h(\tilde{v})] \underbrace{\min\{\mathbb{E}[v_i|\tilde{v}_i = \tilde{v}], M\}}_{=:\beta(\tilde{v})} \, d\tilde{v}$$

$$\geq [H(\tilde{v}|i) - H(\tilde{v})]\beta(\tilde{v})\Big|_{\max\{0,\tilde{v}'\}}^{\infty} - \int_{\max\{0,\tilde{v}'\}}^{\infty} [H(\tilde{v}|i) - H(\tilde{v})] \frac{\partial}{\partial \tilde{v}} \beta(\tilde{v}) \, d\tilde{v}$$

$$= \overbrace{[H(\max\{0,\tilde{v}'\}) - H(\max\{0,\tilde{v}'\}|i)]}^{\geq 0 \text{ by MLRP}} \overbrace{\mathbb{E}[v_i|\tilde{v}_i = \max\{0,\tilde{v}'\}]}^{\geq 0 \text{ by definition of } \tilde{v}'}$$

$$+ \int_{\max\{0,\tilde{v}'\}}^{\infty} \underbrace{[H(\tilde{v}) - H(\tilde{v}|i)]}_{\geq 0 \text{ by MLRP}} \underbrace{\frac{\partial}{\partial \tilde{v}} \beta(\tilde{v})}_{\geq 0 \text{ by MLRP}} d\tilde{v}$$

$$\geq \int_{\max\{0,\tilde{v}'\}}^{\tilde{v}''} [H(\tilde{v}) - H(\tilde{v}|i)] \frac{\partial}{\partial \tilde{v}} \beta(\tilde{v}) \, d\tilde{v}.$$

Using the facts that $H(\tilde{v}) > H(\tilde{v}|i)$ for any $\tilde{v} \leq \tilde{v}''$ and that $\partial \mathbb{E}[v_i|\tilde{v}_i = \tilde{v}]/\partial \tilde{v} > 0$ on any dense subset of $\mathbb{R}$,

$$\lim_{M\to\infty} \int_{\max\{0,\tilde{v}'\}}^{\tilde{v}''} [H(\tilde{v}) - H(\tilde{v}|i)] \frac{\partial}{\partial \tilde{v}} \beta(\tilde{v}) \, d\tilde{v}$$

$$= \int_{\max\{0,\tilde{v}'\}}^{\tilde{v}''} [H(\tilde{v}) - H(\tilde{v}|i)] \frac{\partial}{\partial \tilde{v}} \mathbb{E}[v_i|\tilde{v}_i = \tilde{v}] \, d\tilde{v}$$

$$> 0.$$

'(b) $\Rightarrow$ (a)'. Consider binary steering with a cutoff $\tilde{v}^c \leq 0$. For any realised v_i, m_i with $\tilde{v}(v_i + m_i) \geq 0$, the probability of product i being suggested is independent of v_i and m_i; this probability depends only on the number of other products j with $\tilde{v}_j \geq \tilde{v}^c$. Hence, steering does not change the distribution of v_i and m_i conditional on purchase. As a consequence, if the consumer does not buy reasonably, steering with such a technology decreases the consumer's welfare.

(ii) The proof follows from part (i) because an always reasonable consumer buys reasonably. □

Proof of Proposition 6. (i) Given its knowledge of v_{i^*}, the seller chooses its price to solve

$$\max_p \quad p[1 - F(\bar{w}(p) - v_{i^*})].$$

We show that an optimal price exists for any given value v_{i^*}. Clearly, $p \leq 0$ is suboptimal. The consumer buys at a price $p > 0$ if and only if $w_{i^*} \geq \tilde{v}^{-1}(p) =: \bar{w}(p)$. Because $\tilde{v}(\cdot)$ is strictly increasing and has full range, its inverse is strictly increasing with full range, so that $\lim_{p\to\infty} \bar{w}(p) = \infty$. Furthermore, since $\bar{w}'(p) = 1/\tilde{v}'(\bar{w}(p))$ and, by assumption, $\tilde{v}'(w) \leq 1$ for large enough w, $\lim_{p\to\infty} \bar{w}'(p) > 0$. Hence, since F has an increasing and diverging hazard rate, for large enough p, marginal profits

$$\begin{aligned}
&1 - F(\bar{w}(p) - v_{i^*}) - p\bar{w}'(p)f(\bar{w}(p) - v_{i^*}) \\
&= \bar{w}'(p)f(\bar{w}(p) - v_{i^*})\underbrace{\left[\frac{1}{\bar{w}'(p)}\frac{1 - F(\bar{w}(p) - v_{i^*})}{f(\bar{w}(p) - v_{i^*})} - p\right]}_{<\ 0 \text{ for large enough } p}
\end{aligned}$$

are decreasing. Thus, an optimal price exists, and it satisfies the first-order condition

$$\begin{aligned}
&1 - F(\bar{w}(p) - v_{i^*}) - p\bar{w}'(p)f(\bar{w}(p) - v_{i^*}) = 0 \\
&\text{or} \quad p = \frac{1}{\bar{w}'(p)}\frac{1 - F(\bar{w}(p) - v_{i^*})}{f(\bar{w}(p) - v_{i^*})}.
\end{aligned} \tag{A11}$$

We next study properties of the optimal price $p(v_{i^*})$ as $v_{i^*} \to \infty$. First, $p(v_{i^*})$ cannot converge; if it did, the intermediary's profits would be bounded. But, since $\tilde{v}(\cdot)$ has full range, we can construct a sequence of prices $\hat{p}(v_{i^*})$ such that, for any value v_{i^*}, $\bar{w}(\hat{p}(v_{i^*})) - v_{i^*} \equiv$ constant. Because $\tilde{v}(\cdot)$, and thus also $\bar{w}(\cdot)$, is increasing with full range, this implies that $\hat{p}(v_{i^*}) \to \infty$ as $v_{i^*} \to \infty$. Moreover, since $\bar{w}(\hat{p}(v_{i^*})) - v_{i^*}$ is constant by construction, $\hat{p}(v_{i^*})[1 - F(\bar{w}(\hat{p}(v_{i^*})) - v_{i^*})] \to \infty$ as $v_{i^*} \to \infty$. Hence, the intermediary would have an incentive to deviate to $\hat{p}(v_{i^*})$, a contradiction.

Second, we argue that, as $v_{i^*} \to \infty$, the probability of purchase approaches 1. Because the optimal price approaches infinity as v_{i^*} does, by (A11),

$$\frac{1}{\bar{w}'(p(v_{i^*}))}\frac{1 - F(\bar{w}(p(v_{i^*})) - v_{i^*})}{f(\bar{w}(p(v_{i^*})) - v_{i^*})} \to \infty \quad \text{as } v_{i^*} \to \infty.$$

Recall that $\lim_{p\to\infty} \bar{w}'(p) > 0$. Hence, $f(\bar{w}(p(v_{i^*})) - v_{i^*})/[1 - F(\bar{w}(p(v_{i^*})) - v_{i^*})] \to 0$. Since the hazard rate is increasing, this implies that $\lim_{v_{i^*}\to\infty} \bar{w}(p(v_{i^*})) - v_{i^*} = -\infty$ and $\lim_{v_{i^*}\to\infty} 1 - F(\bar{w}(p(v_{i^*})) - v_{i^*}) = 1$.

Third, we argue that, as $v_{i^*} \to \infty$, $v_{i^*} - p(v_{i^*}) \to \infty$ also; that is, the consumer's surplus becomes arbitrarily large. From above, we know that

$$\begin{aligned}
&\frac{1 - F(\bar{w}(p(v_{i^*})) - v_{i^*})}{f(\bar{w}(p(v_{i^*})) - v_{i^*})} \\
&\quad = \frac{1 - F(\bar{w}(p(v_{i^*})) - p(v_{i^*}) + p(v_{i^*}) - v_{i^*})}{f(\bar{w}(p(v_{i^*})) - p(v_{i^*}) + p(v_{i^*}) - v_{i^*})} \to \infty \quad \text{as } v_{i^*} \to \infty.
\end{aligned} \tag{A12}$$

Moreover, because $\tilde{v}'(w) \leq 1$ for large enough w and $\lim_{p\to\infty} \bar{w}(p) = \infty$ for large enough p,

$$\frac{\partial}{\partial p}\big(\bar{w}(p) - p\big) = \frac{1}{\tilde{v}'(\bar{w}(p))} - 1 \geq 0.$$

Hence, $\lim_{v_{i^*}\to\infty} \bar{w}(p(v_{i^*})) - p(v_{i^*}) > -\infty$. Then (A12) and the increasing hazard rate of F imply that $\lim_{v_{i^*}\to\infty} p(v_{i^*}) - v_{i^*} = -\infty$, so consumer surplus $[v_{i^*} - p(v_{i^*})][1 - F(\bar{w}(\hat{p}(v_{i^*})) - v_{i^*})]$ goes to ∞.

(ii) Because $\tilde{v}(\cdot)$ has full range, we can construct a sequence of prices $p(m_{i^*})$ such that, for any mistake m_{i^*}, $\bar{w}(p(m_{i^*})) - m_{i^*} \equiv$ constant. Because $\tilde{v}(\cdot)$, and thus also $\bar{w}(\cdot)$, is increasing with full range, this implies that $p(m_{i^*}) \to \infty$ as $m_{i^*} \to \infty$. Moreover, because $\bar{w}(p(m_{i^*})) - m_{i^*}$ is constant by construction, $p(m_{i^*})[1 - G(\bar{w}(p(m_{i^*})) - m_{i^*})] \to \infty$ as $m_{i^*} \to \infty$. Hence, the intermediary's profits go to infinity as $m_{i^*} \to \infty$. Because mistake-based steering does not improve the selection of products, total welfare is bounded from above by $\int_0^\infty v \; dG(v) < \infty$. Since total welfare is bounded from above and profits approach infinity, consumer welfare has to approach minus infinity.

(iii) For any realised $\tilde{v}_{i^*} = \tilde{v} > 0$, the seller sets $p(\tilde{v}_{i^*}) = \tilde{v}$. Thus, the consumer's surplus under steering is $\mathbb{E}[v_{i^*} - p(\tilde{v}_{i^*})|\tilde{v}_{i^*} = \tilde{v}] = \mathbb{E}[v_{i^*} - \tilde{v}_{i^*}|\tilde{v}_{i^*} = \tilde{v}] = \mathbb{E}[v_i - \tilde{v}_i|\tilde{v}_i = \tilde{v}]$, where the last equality follows from the fact that, conditional on $\tilde{v}_i$, the intermediary's signals are independent of v_i.

Denote as $\hat{p}$ the price absent steering. The consumer's welfare absent steering is

$$\mathbb{P}[\tilde{v}_i \geq \hat{p}]\mathbb{E}[v_i - \hat{p}|\tilde{v}_i \geq \hat{p}] > \mathbb{P}[\tilde{v}_i \geq \hat{p}]\mathbb{E}[v_i - \hat{p}|\tilde{v}_i = \hat{p}] = \mathbb{P}[\tilde{v}_i \geq \hat{p}]\mathbb{E}[v_i - \tilde{v}_i|\tilde{v}_i = \hat{p}],$$

where the inequality follows from the fact that the consumer's signal structure satisfies MLRP.

By assumption, for sufficiently large perceived values $\tilde{v}$, $\mathbb{E}[v_i - \tilde{v}_i|\tilde{v}_i = \tilde{v}]$ is non-positive. From now on, consider only $\tilde{v}_{i^*} = \tilde{v} > \hat{p}$ for which $\mathbb{E}[v_i - \tilde{v}_i|\tilde{v}_i = \tilde{v}]$ is non-positive. A sufficient condition for steering to decrease the consumer's welfare is then given by

$$\mathbb{E}[v_i - \tilde{v}_i|\tilde{v}_i = \tilde{v}] \leq \mathbb{P}[\tilde{v}_i \geq \hat{p}]\mathbb{E}[v_i - \tilde{v}_i|\tilde{v}_i = \hat{p}].$$

If $\mathbb{E}[v_i - \tilde{v}_i|\tilde{v}_i = \hat{p}] \geq 0$, the above inequality holds. Otherwise, because $\mathbb{E}[v_i - \tilde{v}_i|\tilde{v}_i]$ is weakly decreasing in $\tilde{v}_i$ by assumption and because $\tilde{v} > \hat{p}$, we have

$$\frac{\mathbb{E}[v_i - \tilde{v}_i|\tilde{v}_i = \tilde{v}]}{\mathbb{E}[v_i - \tilde{v}_i|\tilde{v}_i = \hat{p}]} \geq 1.$$

Since $\mathbb{P}[\tilde{v}_i \geq \hat{p}] < 1$, this establishes the claim. □

Heinrich-Heine University, Germany
Central European University, Austria
Institute on Behavior & Inequality (briq), Germany

References

ACCC. (2019). 'Digital platform inquiry: Final report', Australian Competition & Consumer Commission, https://www.accc.gov.au/system/files/Digital%20platforms%20inquiry%20-%20final%20report.pdf.

Acemoglu, D., Makhdoumi, A., Malekian, A. and Ozdaglar, A. (2022). 'Too much data: Prices and inefficiencies in data markets', *American Economic Journal: Microeconomics*, vol. 14(4), pp. 218–56.

Agarwal, S., Amromin, G., Ben-David, I. and Evanoff, D.D. (2016). 'Loan product steering in mortgage markets', Working paper, National Bureau of Economic Research.

Anagol, S., Cole, S. and Sarkar, S. (2017). 'Understanding the advice of commissions motivated agents: Theory and evidence from the Indian life insurance market', *Review of Economics and Statistics*, vol. 99(1), pp. 1–15.

Bagnoli, M. and Bergstrom, T. (2005). 'Log-concave probability and its applications', *Economic Theory*, vol. 26, pp. 445–69.

Bar-Gill, O. (2021). 'Price discrimination with consumer misperception', *Applied Economics Letters*, vol. 28(10), pp. 829–34.

Barlow, R.E., Marshall, A.W. and Proschan, F. (1963). 'Properties of probability distributions with monotone hazard rate', *The Annals of Mathematical Statistics*, vol. 34(2), pp. 375–89.

Bashir, M.A., Farooq, U., Shahid, M., Zaffar, M.F. and Wilson, C. (2019). 'Quantity vs. quality: Evaluating user interest profiles using ad preference managers', in *Proceedings 2019 Network and Distributed System Security Symposium*, San Diego, CA, 24–27 February.

Bergemann, D. and Bonatti, A. (2011). 'Targeting in advertising markets: Implications for offline versus online media', *RAND Journal of Economics*, vol. 42(3), pp. 417–43.

Bergemann, D., Bonatti, A. and Gan, T. (2022). 'The economics of social data', *RAND Journal of Economics*, vol. 53(2), pp. 263–96.

Bordalo, P., Gennaioli, N. and Shleifer, A. (2013). 'Salience and consumer choice', *Journal of Political Economy*, vol. 121(5), pp. 803–43.
Bordalo, P., Gennaioli, N. and Shleifer, A. (2015). 'Competition for attention', *Review of Economic Studies*, vol. 83(2), pp. 481–513.
Busse, M.R., Pope, D.G., Pope, J.C. and Silva-Risso, J. (2015). 'The psychological effect of weather on car purchases', *Quarterly Journal of Economics*, vol. 130(1), pp. 371–414.
Chang, T.Y., Huang, W. and Wang, Y. (2018). 'Something in the air: Pollution and the demand for health insurance', *Review of Economic Studies*, vol. 85(3), pp. 1609–34.
CMA. (2020). 'Online platforms and digital advertising: Market study final report', Competition and Markets Authority, https://assets.publishing.service.gov.uk/media/5efc57ed3a6f4023d242ed56/Final_report_1_July_2020_.pdf.
Conlin, M., O'Donoghue, T. and Vogelsang, T.J. (2007). 'Projection bias in catalog orders', *American Economic Review*, vol. 97(4), pp. 1217–49.
Crémer, J., de Montjoye, Y.A. and Schweitzer, H. (2019). 'Competition policy for the digital era final report', European Commission, https://ec.europa.eu/competition/publications/reports/kd0419345enn.pdf.
Datta, A., Tschantz, M.C. and Datta, A. (2015). 'Automated experiments on ad privacy settings: A tale of opacity, choice, and discrimination', *Proceedings on Privacy Enhancing Technologies 2015*, vol. 2015(1), pp. 92–112.
de Corniere, A. (2016). 'Search advertising', *American Economic Journal: Microeconomics*, vol. 8(3), pp. 156–88.
de Corniere, A. and de Nijs, R. (2016). 'Online advertising and privacy', *RAND Journal of Economics*, vol. 47(1), pp. 48–72.
de Corniere, A. and Taylor, G. (2019). 'A model of biased intermediation', *RAND Journal of Economics*, vol. 50(4), pp. 854–82.
de los Santos, B. (2018). 'Consumer search on the internet', *International Journal of Industrial Organization*, vol. 58, pp. 66–105.
Eliaz, K. and Spiegler, R. (2006). 'Contracting with diversely naive agents', *Review of Economic Studies*, vol. 73(3), pp. 689–714.
Eliaz, K. and Spiegler, R. (2016). 'Search design and broad matching', *American Economic Review*, vol. 106(3), pp. 563–86.
European Commission. (2020). 'Proposal for a regulation of the European Parliament and the council on a single market for digital services (digital services act) and amending directive 2000/31/ec', https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM:2020:825:FIN.
Fairfield, J.A.T. and Engel, C. (2015). 'Privacy as a public good', *Duke LJ*, vol. 65(3), p. 385.
Fletcher, A., Crawford, G.S., Crémer, J., Dinielli, D., Heidhues, P., Luca, M., Salz, T., Schnitzer, M., Morton, F.M.S., Seim, K. and Sinkinson, M. (2023). 'Consumer protection for online markets and large digital platforms', *Yale Journal on Regulation*, forthcoming.
Furman, J., Coyle, D., Fletcher, A., McAuley, D. and Marsden, P. (2019). 'Unlocking digital competition: Report of the digital competition expert panel', UK Government Publication, HM Treasury, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf.
Gabaix, X. and Laibson, D. (2006). 'Shrouded attributes, consumer myopia, and information suppression in competitive markets', *Quarterly Journal of Economics*, vol. 121(2), pp. 505–40.
Gottlieb, D. (2012). 'Prospect theory, life insurance, and annuities', Working paper, The Wharton School, University of Pennsylvania Research Paper Series No 44.
Hagiu, A. and Jullien, B. (2011). 'Why do intermediaries divert search?', *RAND Journal of Economics*, vol. 42(2), pp. 337–62.
Hannak, A., Soeller, G., Lazer, D., Mislove, A. and Wilson, C. (2014). 'Measuring price discrimination and steering on e-commerce web sites', in *Proceedings of the 2014 Conference on Internet Measurement Conference*, pp. 305–18, New York: Association for Computing Machinery.
Hansen, A. (2020). 'The three extreme value distributions: An introductory review', *Frontiers in Physics*, vol. 8, p. 533.
Heidhues, P. and Kőszegi, B. (2010). 'Exploiting naivete about self-control in the credit market', *American Economic Review*, vol. 100(5), pp. 2279–303.
Heidhues, P. and Kőszegi, B. (2017). 'Naivete-based discrimination', *Quarterly Journal of Economics*, vol. 132(2), pp. 1019–54.
Heidhues, P. and Kőszegi, B. (2018). 'Behavioral industrial organization', in (B.D. Bernheim, S. DellaVigna and D.I. Laibson, eds.), *Handbook of Behavioral Economics: Applications and Foundations 1*, pp. 517–612, Amsterdam: North-Holland.
Hermalin, B. and Katz, M. (2006). 'Privacy, property rights and efficiency: The economics of privacy as secrecy', *Quantitative Marketing and Economics*, vol. 4(3), pp. 209–39.
Hidir, S. and Vellodi, N. (2021). 'Privacy, personalization, and price discrimination', *Journal of the European Economic Association*, vol. 19(2), pp. 1342–63.
Inderst, R. and Ottaviani, M. (2012a). 'Competition through commissions and kickbacks', *American Economic Review*, vol. 102(2), pp. 780–809.
Inderst, R. and Ottaviani, M. (2012b). 'Financial advice', *Journal of Economic Literature*, vol. 50(2), pp. 494–512.
Johnen, J. (2020). 'Dynamic competition in deceptive markets', *RAND Journal of Economics*, vol. 51(2), pp. 375–401.

Kohavi, R., Tang, D. and Xu, Y. (2020). *Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing*, Cambridge: Cambridge University Press.
Kőszegi, B. and Szeidl, A. (2013). 'A model of focusing in economic choice', *Quarterly Journal of Economics*, vol. 128(1), pp. 53–107.
Koulayev, S. (2014). 'Search for differentiated products: Identification and estimation', *RAND Journal of Economics*, vol. 45(3), pp. 553–75.
Loewenstein, G., O'Donoghue, T. and Rabin, M. (2003). 'Projection bias in predicting future utility', *Quarterly Journal of Economics*, vol. 118(4), pp. 81–123.
MacCarthy, M. (2010). 'New directions in privacy: Disclosure, unfairness and externalities', *ISJLP*, vol. 6(3), p. 425.
Marotta, V., Zhang, K. and Acquisti, A. (2018). 'The welfare impact of targeted advertising', Working paper, Carnegie Mellon University.
Mathur, A., Acar, G., Friedman, M.J., Lucherini, E., Mayer, J., Chetty, M. and Narayanan, A. (2019). 'Dark patterns at scale: Findings from a crawl of 11K shopping websites', *Proceedings of the ACM on Human-Computer Interaction*, vol. 3(CSCW), pp. 1–32.
McMahan, H.B., Holt, G., Sculley, D., Young, M., Ebner, D., Grady, J., Nie, L., Phillips, T., Davydov, E., Golovin, D., Chikkerur, S., Liu, D., Wattenberg, M., Hrafnkelsson, A.M., Boulos, T. and Kubica, J. (2013). 'Ad click prediction: A view from the trenches', in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1222–30, New York: Association for Computing Machinery.
Milgrom, P.R. (1981). 'Good news and bad news: Representation theorems and applications', *The Bell Journal of Economics*, vol. 12(2), pp. 380–91.
Monopolkommission. (2020). *XXIII. Hauptgutachten der Monopolkommission*, https://monopolkommission.de/images/HG23/HGXXIII_Gesamt.pdf, Baden-Baden: Nomos.
Mullainathan, S., Nöth, M. and Schoar, A. (2011). 'The market for financial advice: An audit study', Working paper, Massachusetts Institute of Technology.
Murooka, T. (2015). 'Deception under competitive intermediation', Working paper, University of Munich.
Posner, R.A. (1981). 'The economics of privacy', *American Economic Review*, vol. 71(2), pp. 405–9.
Rabin, M. and Vayanos, D. (2010). 'The gambler's and hot-hand fallacies: Theory and applications', *Review of Economic Studies*, vol. 77(2), pp. 730–78.
Schwartzstein, J. and Sunderam, A. (2021). 'Using models to persuade', *American Economic Review*, vol. 111(1), pp. 276–323.
Simonsohn, U. (2010). 'Weather to go to college', Economic Journal, vol. 120(543), pp. 270–80.
Stigler, G.J. (1980). 'An introduction to privacy in economics and politics', *Journal of Legal Studies*, vol. 9(4), pp. 623–44.
Teh, T.-H. and Wright, J. (2022). 'Intermediation and steering: Competition in prices and commissions', *American Economic Journal: Microeconomics*, vol. 14(2), pp. 281–321.
Varian, H.R. (1996). 'Economic aspects of personal privacy', Working paper, UC Berkeley.
Vincent, M. (2020) 'UK regulator says Google not doing enough about scam ads', *Financial Times*, September 24.
Wills, C.E. and Tatar, C. (2012). 'Understanding what they do with what they know', in *Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society*, pp. 13–18, New York: Association for Computing Machinery.