Original Research Article

Big Data & Society
July–December: 1–15

DOI: 10.1177/20539517221132535
journals.sagepub.com/home/bds

The Concentration-after-Personalisation Index (CAPI): Governing effects of personalisation using the example of targeted online advertising

Johann Laux[1], Fabian Stephany[1,2], Chris Russell[1], Sandra Wachter[1] and Brent Mittelstadt[1]

Abstract

Firms are increasingly personalising their offers and services, leading to an ever finer-grained segmentation of consumers online. Targeted online advertising and online price discrimination are salient examples of this development. While personalisation's overall effects on consumer welfare are expectably ambiguous, it can lead to concentration in the distribution of advertising and commercial offers. Constellations are possible in which a market is generally open to competition, but the targeted consumer is only made aware of one possible seller. For the consumer, such a market could effectively resemble a monopoly. We call such extreme cases 'targeting pockets'. Competition-law metrics such as the Herfindahl–Hirschman Index and traditional means of public oversight of adverts would not detect this concentration. We, therefore, suggest a novel metric, the Concentration-after-Personalisation Index (CAPI). The CAPI treats every consumer as a separate 'market', computes a measure of concentration for personalised adverts and offers for each individual consumer separately, and then averages the result to measure the exposure experienced by an average consumer. We demonstrate how the CAPI can serve as a monitoring tool for regulators and auditors and thus help to enforce existing consumer law as well as proposed new regulations such as the European Union's Digital Services Act and its Artificial Intelligence Act. We further show how adding noise via randomly distributed non-personalised adverts can dilute the potential harm of overly concentrated personalisation. We demonstrate how the CAPI can identify the optimal degree of added noise, balancing the protection of consumer choice with the economic interests of advertisers.

Keywords

Targeted advertising, personalisation, consumer welfare, consumer protection, consumer law, competition law, EU law, platform regulation, Digital Services Act, AI Act, Unfair Commercial Practices Directive, digital markets, law and economics, novel metrics

Introduction

Today's technology enables firms to personalise their interaction with consumers to an unprecedented degree. Personalisation of online advertising, for example, promises to improve consumer engagement with adverts by providing information predicted to be more relevant to the individual consumer (Aguirre et al., 2016). Through 'behavioural targeting' advertisers can track a consumer's online behaviour such as regularly visiting a website for runners and liking social media posts by athletes (see further Laux et al., 2021a; Wachter, 2021). Advertisers can then infer that this consumer is a good target for adverts promoting running shoes. Information collected online also allows sellers to tailor prices according to the predicted price sensitivity of a consumer. They can then differentiate the online price for identical products or services, engaging in what is called 'price discrimination' (Bar-Gill, 2019; Poort and Zuiderveen Borgesius, 2021; Strahilevitz, 2013). They may charge higher prices for returning customers to a

[1]Oxford Internet Institute, Oxford, UK
[2]Humboldt Institute for Internet and Society, Berlin, Germany

Corresponding author:
Johann Laux, Oxford Internet Institute, 1 St Giles, Oxford, OX1 3JS, UK.
Email: johann.laux@oii.ox.ac.uk

given online store or to consumers without a physical shoe store in their neighbourhood (Poort and Zuiderveen Borgesius, 2021: 2).

Such personalisation of both advertising and pricing affects – amongst other normative concerns (see further Hacker, 2021: 1–2; Laux et al., 2021a: 726–29) – the welfare of consumers and sellers, or 'producers' (Ezrachi and Stucke, 2016a: 586). Whether or not the overall welfare effect is positive or negative depends on the relative strength of the effect for consumers and producers (Marotta et al., 2021). Both can experience positive and negative welfare effects from personalisation (Marotta et al., 2021; Poort and Zuiderveen Borgesius, 2021). A key determinant for the overall welfare effect is the granularity of personalisation: the smaller and more homogenous the segmentation of consumers is, the more welfare is likely to be shifted from consumers to producers (Poort and Zuiderveen Borgesius, 2021: 5). However, from a normative perspective, adverts and offers which are distributed to 'out-of-target' audiences are not without value.

The economic incentive to increase personalisation raises the question of how much segmented consumers are aware of offers in the market beyond those which they are being targeted with (see also Laux et al., 2021a). Constellations are possible in which a market is generally open to competition, but the targeted consumer is only made aware of one possible seller. From the consumer's perspective, such a market could effectively resemble a monopoly. Likewise, a given seller may target certain consumers exclusively with one variant of its product or service. We do not expect such extreme 'targeting pockets' to be the most common effect of personalisation. Even with targeted adverts, many consumers will regularly be exposed to offers made by more than one seller. Most consumers do not receive advertising exclusively online and, at least in urban areas, there are usually options to visit a physical shop. However, consumers' interaction with the market proceeds to an unprecedented degree online (Calo, 2014: 1003–1004). In this context, the use of automated bandit-based bidding systems in targeted ads is particularly concerning. Interactions between such algorithmic systems have been shown to result in collusion-like behaviour and overpricing (Calvano et al., 2020: 3267; Hansen et al., 2021: 2; for an empirical investigation, see Assad et al., 2020: 5). Moreover, many products and services are exclusively offered online – think of online gaming, for example. The risk for consumers of being exposed to offers by just one seller can thus neither be ruled out theoretically nor are the potential negative economic consequences for the consumer negligible. In situations where more than one seller remains, a high degree of personalisation by one seller could still extract more of the consumer surplus than without such customisation and have negative welfare effects (Acemoglu, 2021: 11–14).

Finding a measure to screen personalised adverts and offers for heightened concentration could thus help to reduce the potential for economic harm to consumers. It would also allow to test whether (and if so, reveal where) extreme cases of 'targeting pockets' occur. To date, we are not aware of an existing measure that implements this type of screening. We, therefore, suggest a novel metric, the Concentration-after-Personalisation Index (CAPI). The CAPI builds on a mathematical index used in competition law, the Herfindahl–Hirschman Index (HHI). While the HHI is a measure of concentration in the market, the CAPI develops it further and treats every consumer as a separate 'market'. The formula presented in this paper computes a measure analogous to the HHI for each consumer to measure concentration in their exposure to personalised offers and services.

Imagine that there are two neighbouring towns with one shoe store in each, but no one travels between these towns. Each town could then be seen as being effectively in a monopoly even though the HHI computed for the two towns together would not reveal a monopoly market. As regards online advertising, although there may be more than one seller of a given product or service, due to targeting, consumers may not be served with advertising from more than one seller or be served with adverts from significantly fewer sellers than without targeting. The CAPI offers a metric to measure such concentration at the level of the individual consumer. As will be shown in the section 'A novel metric: the CAPI', we generate the CAPI by computing the HHI in terms of advert exposure for each individual separately and then average the individual HHI scores. This gives a measure of the diversity of adverts experienced by an average individual.

The CAPI is intended to function as a tool with which regulators and auditors can screen personalisation measures such as online adverts for low degrees of diversification amongst sellers for a typical consumer. A high score on the CAPI provides statistical evidence for a heightened risk of consumer harm due to low diversification. It cannot, however, replace the contextual judicial interpretation as to whether a low diversification amongst sellers in online adverts amounts to a breach of law. CAPI screening can support the enforcement of regulatory law in the context of algorithmic targeting and personalisation, including novel legislative proposals for platform regulation and transparency in online advertising. There is great value in discovering 'targeting pockets' via the CAPI, not least because public oversight of advertising is compromised due to the fragmentation of consumers by targeted adverts, as we show below. Applying the CAPI can further help regulators to direct regulatory resources to areas in the distribution of adverts where the index shows the risk of consumer harm to be heightened. Moreover, the CAPI offers a tool with which to analyse the data captured in advertising repositories suggested in recent regulatory proposals such as the European Union's (EU) Digital Services Act (DSA) and new transparency rules on political advertising and targeting presented by the European Commission (2020b, 2021c). While a novel metric, the CAPI is thus a fairly 'conservative' measure as it can be implemented within the current (and proposed) regulatory framework.

While this paper draws on the example of advertising, CAPI screening could be applied in various contexts of personalisation. Recommender systems are one such area. As the development of Artificial Intelligence (AI) is currently on a trajectory to increase personalisation of offers and products, CAPI screening could prove to become useful as an auditing tool in AI governance. The customisation of products may allow AI-intensive firms to capture more of the consumer surplus and lead to higher prices due to a lack of price competition in the market (Acemoglu, 2021). The European Commission not least proposed its AI Act with a view to 'facilitating audits of AI systems' (European Commission, 2021a: 10).

This paper proceeds as follows: we begin by describing some of the risks to consumer welfare and law enforcement which personalisation via targeted advertising and personalised pricing can pose (see section 'Personalisation, consumer welfare and law enforcement'). In the EU single market order, the consumer welfare standard seeks to safeguard competitive markets, protect consumer choice, and prevent higher price levels and/or lower quality than would have otherwise prevailed (European Commission, 2009: para 19; Ezrachi, 2018: 4–7; Weatherill, 2016). While new forms of consumer vulnerability due to behavioural targeting and price discrimination in the digital economy have increasingly been anticipated by the literature, current EU consumer protection law does not reflect this development and adheres to a widely rationalistic consumer standard (Helberger et al., 2021; Laux et al., 2021a).

We then show how we construct the CAPI based on the HHI and how we envision its usage (see section 'A novel metric: the CAPI'). Although there are legal proposals to build advertising repositories, access to data on targeted advertising and its distribution is currently scarce. We, therefore, compute the CAPI by drawing on synthetic data to provide sample calculations based on a fictitious scenario of an online market for mattresses. When calculating the CAPI, outcomes are possible in which there is a large gap between the CAPI and the HHI. This means that the diversity of adverts delivered (as measured by the HHI) does not match the diversity of adverts that the average individual is exposed to (as measured by the CAPI). We draw on the CAPI's variance to reveal that there is a group of consumers which experiences a distinctly different, concentrated exposure to adverts than other consumers. We show that by deploying a particularly strong form of targeting, one seller can trap consumers in a 'targeting pocket' without needing significant market power to do so.

A gap between the computed scores provides a motivation for regulatory intervention, as market concentration metrics such as the HHI cannot spot the lack of diversity individual consumers are experiencing in their personalised interaction with the market. We, therefore, continue by suggesting a regulatory solution based on the CAPI: the adding of optimised degrees of noise to the targeting, thus serving consumers with offers that are not based on their consumer profile (see section 'A novel solution: calculating the optimal degree of added noise with the CAPI'). Currently, the EU appears to be enacting a partial ban of targeted advertising in its DSA, prohibiting targeting based on data of minors and limiting targeting variables through excluding categories of sensitive data such as race, religion, or political orientation (European Parliament, 2022). At the same time, online advertising providers and intermediaries are rolling out second-generation tracking technology, meant to safeguard privacy by limiting variables used for targeting (see sections 'Persuasive advertising and rational consumer choice' and 'Regulatory options and second-generation tracking'). Adding noise, however, acknowledges the ambiguous consumer welfare effects of personalisation laid out in the section 'Personalisation, consumer welfare and law enforcement'. A complete ban of all personalisation would erase all potential welfare benefits (see sections 'Persuasive advertising and rational consumer choice' and 'Regulatory options and second-generation tracking'). Other normative considerations such as privacy may, of course, favour a ban of all targeting, but they must remain outside the scope of this paper. Based on the CAPI, we show that there is a computable optimal degree of added noise, which balances the normative goals of protecting consumer choice and the economic interests of advertisers. Adding noise with the CAPI can thus present a proportionate response to the risks of consumer harm from personalisation.

We conclude by stating the need for better access to data for future research, differentiating targeting pockets from 'filter bubbles', and discussing the contribution CAPI could make to prevent harms from personalisation beyond the domain of targeted advertising (see section 'Conclusion: personalisation beyond advertising').

Personalisation, consumer welfare and law enforcement

The economic incentive for sellers to segment consumers into ever-smaller groups invites reflecting on the risks for consumers involved. Below, we consider several (negative) effects of personalisation on consumer welfare, including distributional consequences from shifting surplus from consumers to sellers through behavioural targeting and personalised pricing. We also consider the (negative) effect of personalisation on public oversight of adverts and the resulting consequences for law enforcement. The aim of this section is to outline our theory of harm behind concentration in personalisation which can motivate regulatory intervention supported by the CAPI. Positive effects of personalisation for consumers, such as a potential reduction of choice overload (André et al., 2018: 32), will thus find comparatively less consideration below.

While we do not address privacy harms directly in this paper, consumers' privacy has links to their welfare. Consumers may value privacy instrumentally, as it allows them to enjoy a larger consumer surplus if sellers know less about them and cannot charge higher prices. Consumers may also value privacy intrinsically, not least to avoid targeted ads (Acemoglu, 2021: 6). While some have argued that rational consumers would only consent to having companies use their data if the anticipated benefits exceed the costs to their privacy, this view has been challenged, not least due to the social dimension of data (Acemoglu, 2021: 6). An individual sharing their data also provides information about others, thus creating externalities (Acemoglu, 2021: 6).

Below, we are going to consider the effects of personalisation on diversity (see section 'Diversity in consumer choice') and rationality (see section 'Persuasive advertising and rational consumer choice') in consumer choice, the distributional consequences of price discrimination (see section 'Price discrimination'), and the effects of segmenting advertising audiences on law enforcement (see section 'Public oversight and law enforcement').

Diversity in consumer choice

As mentioned, the EU single market order imposes a consumer welfare standard that protects competition as well as consumer choice. By segmenting its audience, personalisation can substantially reduce the range of choice consumers experience, for example, in the advertising they receive (Laux et al., 2021a).

In the extreme case, an individual consumer may be targeted by one seller with one product or service only. For such a 'targeting pocket' to occur, it would not be necessary for a single firm to have gained significant market power, as we will show in the section 'A novel metric: the CAPI'. The HHI would therefore not detect concentration in the market, and we would not speak of an actual 'monopoly'. The extreme targeting case is not negligible, as from the perspective of the individual consumer, it may have similar effects in terms of choice, pricing, and quality as a monopoly. Consider products and services that are only offered online, such as online games; without advertisement, their existence may be almost undiscoverable for consumers. A single game developer may offer more than one game and yet still target individuals with one game from their portfolio of games only.

In less extreme constellations, consumers may only be targeted with a small subset of available products and services. For offers that are available online and offline, this may not be too different from long-standing practices of advertising, in which large marketing budgets have allowed certain sellers to outbid their competitors for advertising space. The difference with online targeted advertising, however, is the degree of consumer segmentation that is both technically possible and economically affordable for many, if not most, sellers (Goldfarb, 2014). As we show in the section 'A novel metric: the CAPI', a seller with a comparatively small advertising budget can still place consumers in a targeting pocket.

For the EU single market with its emphasis on safeguarding consumer choice, the effects of personalisation, including the risk of 'targeting pockets', must therefore be of normative concern.

Persuasive advertising and rational consumer choice

While consumers may benefit from a better matching of ads with their preferences, behaviourally targeted advertising can also exploit vulnerabilities in their ability to take rational decisions. Today's ad-technology can predict consumer behaviour and biases beyond what consumers themselves can be expected to know or understand (Acemoglu, 2021: 6). Consequently, consumers may be persuaded to buy products they did not intend to buy (Laux et al., 2021a: 726–29; see also Hacker, 2021; Strahilevitz, 2013; Zuiderveen Borgesius, 2015: 15, 47). Such behavioural manipulation can shift surplus from the consumer to the seller (and may distort the composition of consumption, thus creating new inefficiencies, cf. Acemoglu, 2021: 18). Again, this threatens the normative aim of the EU single market order to safeguard certain choice conditions in which (rational) consumers are able to take rational economic decisions.

Take psycho-graphic targeting as an example: it allows to determine the current emotional state of consumers and predict how susceptible they are to advertising based on their mood (Matz and Netzer, 2017). Another example of companies estimating 'prime vulnerability moments' is the anecdotal evidence of retailers forecasting whether women are pregnant and sending them hidden ads for baby products (Acemoglu, 2021: 14). While consumers may have learned to protect themselves against abusive forms of persuasive advertising they have been exposed to since the very beginning of advertising, Big Data analytics and machine learning technology may have quickly outdated consumers' learnings (Acemoglu, 2021: 14).

Advertising arguably always had some persuasive effect, altering consumers' preferences and creating artificial product differentiation (Bagwell, 2007: 1705–1706; for a discussion related to targeted advertising, see Laux et al., 2021a: 724–26). If advertising was purely persuasive and not informative, it would offer no real value to consumers and likely have anti-competitive effects. Seen from this extreme, adverts manipulate consumers into buying products they did not intend to buy. They harm competitors who sell the products that consumers would otherwise have bought (Laux et al., 2021a: 724–5; Woodcock, 2018: 2272). Advertising thus has a negative social value. Vice versa, if advertising was purely informative and not persuasive, it would have pro-competitive effects as consumers receive valuable information about products' existence, their price, and their quality in markets characterised by imperfect consumer information and high search costs (Bagwell, 2007: 1705–1706; see also Laux et al., 2021a: 724–6). Adverts then promote competition between established firms and facilitate market entry for new competitors (Woodcock, 2018: 2272–2273). Neither extreme view on the social value of advertising appears to be entirely convincing.[1]

Seen from a purely informative view on advertising, one may argue that as much as targeting actually increases the matching of adverts with consumers' preferences, it can amplify advertising's informative effect. This will then regularly depend on the quality of the data and the inferential analytics deployed in targeting. Putting questions about the remaining value of information about products outside of the predicted range of preferences aside, rational consumers should thus prefer to be tracked and targeted online. Actual consumer behaviour, however, tells a different story. Consumers tend to avoid being tracked for advertising purposes (Ham, 2017). They now take advantage of changes to mobile phone operating systems by Apple and Google in 2021 that allow users to opt out of tracking for advertising by apps (Grant, 2021; Meaker, 2022). In response, advertisers cut back their spending on apps such as Snap, Twitter, Facebook and YouTube as the accuracy of targeting allegedly decreased (Filippino et al., 2021; McGee, 2021).

Consumers' ability to opt out of mobile-app tracking does not mean the end of the data-based personalisation of online advertising and commercial communication. Instead, online advertisers are developing a new generation of tracking tools that allow the delivery of personalised ads without the use of third-party cookies (Meaker, 2022; for a critique, see Cyphers, 2021; Milmo, 2022). Consumers thus remain exposed to risks stemming from behavioural targeting, not to speak of those consumers who decide to opt in and allow tracking by mobile apps.

Price discrimination

As stated above, price discrimination can have ambiguous effects on consumer welfare. It can lead to an expansion of supply as more consumers can be served (Bourreau et al., 2017: 53; Marotta et al., 2021). Sellers can offer their products and services to consumers with smaller budgets (who thus benefit), which increases total consumption. At the same time, this will regularly lead to higher prices for consumers with a (predicted) higher willingness to pay. For these types, consumer surplus – i.e., the difference between the price a consumer pays and the price she would be willing to pay (Ezrachi and Stucke, 2016b: 86) – is reduced. Especially those consumers who have limited choice of whom to buy from or with little motivation to 'shop around' will likely face discrimination (ACCC, 2019: 517; Strahilevitz, 2013).

Furthermore, price discrimination not only has ambiguous welfare effects for consumers but potentially also for producers: while personalisation allows sellers to efficiently allocate their advertising costs, they may suffer a net welfare loss if competition for price-sensitive consumers is intensified through price discrimination (Poort and Zuiderveen Borgesius, 2021: 5).

Public oversight and law enforcement

Risks not only occur at the individual level of consumers but also for public law enforcement. While a billboard can be seen by all passers-by, targeted ads are only visible to the targeted audience. Public oversight of personalised adverts can thus be compromised. This is problematic, as the targeted groups may often lack the contextual knowledge of why they are being targeted and why it can be harmful to them (Milano et al., 2021; Wachter, 2021). Members of targeted groups cannot always be expected to understand their susceptibility to the targeted messages, and those outside of such groups are less likely to learn about the targeting and pursue complaints on behalf of the targeted. Complaint-based systems of consumer protection are therefore threatened in their efficacy (Wachter, 2021).

A novel metric: The CAPI

In recent years, concerns about the possible negative effects of personalised online services have led to efforts to quantify their degree of personalisation. For example, personalisation in web search as well as online news has been measured by using similarity ratios (Cozza et al., 2016; Dos Santos et al., 2020; Hannák et al., 2017; Krafft et al., 2019; Le et al., 2019; Puschmann, 2018; Salehi et al., 2015). Similarity ratios have also been applied to measuring targeted online advertising (Balebako et al., 2012; Guha et al., 2010). As mentioned, the idea behind the CAPI is instead to treat every consumer as a 'market' and measure the concentration of adverts a typical individual receives from one particular shop type. Building on the HHI, the CAPI measures 'market' concentration in terms of exposure to adverts rather than sales.

This section begins by addressing the obvious need to feed the CAPI with data if it was to be implemented. The accessibility of personalisation data, however, is a significant problem (see section 'Preliminary consideration: access to data'). Fortunately, EU policymakers are currently drafting legislation that mandates online advertising repositories. If enacted, the CAPI could help to implement and enforce the new legislation. We then show how the CAPI is created in three consecutive steps. First, we adapt the HHI by calculating the index not for market shares (via sales) but for exposure to adverts (see section 'Step 1: adapting the HHI to measure ad exposure'). Second, we arrive at the CAPI by computing the HHI in terms of advert exposure for each individual separately and then average the result to measure the diversity of adverts experienced by an average individual (see section 'Step 2: arriving at the CAPI'). Third, we show how to work with the CAPI and what its scores can tell us about risks to consumers (see section 'Step 3: working with the CAPI'). We circumnavigate the problem of lacking data by drawing on a fabricated example: an online market for mattresses. Up to this point, we calculate the CAPI and analyse its results without knowledge about targeted groups. We, therefore, show an additional calculation, assuming knowledge about which groups of consumers have been targeted (see section 'Consequent consideration: lifting the veil of group membership'). This final point is an important preliminary step for the regulatory solution to 'targeting pockets' suggested in the following section.

Preliminary consideration: Access to data

When implemented in practice, the CAPI would be computed based on actual data. As regards our example of targeted adverts, while available to advertisers and ad-tech providers, such data is currently difficult to access for researchers and public authorities (Laux et al., 2021b). Encouragingly, recent legislative attempts in the EU to introduce audits and risk assessments to targeted advertising are also aiming to improve data accessibility.

As mentioned, the DSA envisions the creation of advertising repositories for very large online platforms in Article 30(2)(e) DSA. Likewise, the proposed legislation on political advertising by the European Commission potentially requires record-keeping of (targeted) political adverts by online intermediaries and advertising service providers (European Commission, 2020a: 5). Moreover, the proposed AI Act of the EU seeks to guarantee that, amongst other actors, auditors and researchers 'should be able to access and use high quality datasets' (European Commission, 2021a: Recital 45). After all, AI is increasingly the technological driver behind targeted advertising and personalised pricing strategies (Gonzales, 2022).

Depending on their final wording and especially of Article 30(2) DSA, these proposals could provide enough data to calculate CAPI scores for advertisements displayed on large platforms in the EU. The CAPI could thus prove to be a helpful tool for monitoring consumer harms and 'targeting pockets' as described above, facilitating the implementation of these novel regulations as well as enhancing the enforcement of existing consumer law such as the Unfair Commercial Practices Directive (UCPD) and, potentially, competition law.

Step 1: Adapting the HHI to measure ad exposure

As mentioned, we are creating the CAPI based on the HHI. The latter is used in competition law enforcement as a measure of concentration in the market (White, 2013: 39–40). The HHI is calculated by squaring the market share of each firm in the market and then taking the sum of the resulting squares (United States Department of Justice, 2018). A sample calculation for a market of four firms with shares of 30, 30, 20 and 20% goes as follows: $0.3^2 + 0.3^2 + 0.2^2 + 0.2^2 = 0.26$. Here, we are using an HHI that describes market shares as fractions of 1. Therefore, the maximal number that our HHI could reach is 1, in case one firm is the only player in the market. It approaches 0 when a market is shared by a large number of firms of relatively equal size. Two factors let the HHI increase: when the number of firms in the market decreases and when the disparity in size between those firms increases (United States Department of Justice, 2018).[2]

In this paper, we adapt the HHI to create the CAPI to measure exposure to adverts of a given product or service type instead of firms in the market. The idea behind the CAPI is to quantify how consumers perceive the online market as a whole, led by concerns about consumer choice and law enforcement articulated in the section 'Personalisation, consumer welfare and law enforcement'. To achieve this, we need to compute the measure at the level of the individual consumer, as the next step is going to show.

Step 2: Arriving at the CAPI

We suggest the CAPI as a metric in which we compute the HHI in terms of advert exposure for each individual separately and then average the result to measure the diversity of adverts experienced by an average individual.

$$CAPI(x) = \frac{1}{n}\sum_{i=0}^{n} HHI(x_i)$$

where n is the number of individuals, $HII(x_i)$ is the advert exposure of the ith individual computed as follows:

$$HHI(x_i) = \frac{1}{n}\sum_{j=0}^{m}\left(\frac{a_{i,j}}{\sum_{k=0}^{m} a_{i,k}}\right)^2$$

Here m is the total number of advertisers, and $a_{i,j}$ measures the total number of times individual i sees an advertisement from advertiser j.

The CAPI analyses one side effect of 'targeting pockets', i.e., the (lack of) diversity in adverts experienced by an average individual. It does not provide a measure for the degree of personalisation of just one seller's advert. Even if a seller would hyper-personalise ads and address an individual consumer with adverts, as long as that consumer would also be exposed to adverts by competing sellers, the CAPI would not register the hyper-personalisation done by one seller.

We consider online advertising for mattresses as an example. Imagine that one online seller of mattresses buys up all online advertising space for mattresses. Each individual consumer thus receives adverts only from this one seller, and the HHI and CAPI would both equal 1.

Now assume that there are three mattress stores all advertising to approximately one-third of the population, but each individual only receives adverts from one store. Then, for each individual x_i:

$$HHI(x_i) = \frac{1}{n}\sum_{j=0}^{m}\left(\frac{a_{i,j}}{\sum_{k=0}^{m} a_{i,k}}\right)^2 = 1$$

and

$$CAPI(x) = \frac{1}{n}\sum_{i=0}^{n} HHI(x_i) = 1$$

On the other hand, if each individual receives a uniform mix of adverts from across all stores, we find:

$$HHI(x_i) = \frac{1}{n}\sum_{j=0}^{m}\left(\frac{a_{i,j}}{\sum_{k=0}^{m} a_{i,k}}\right)^2 = 0.33$$

and

$$CAPI(x) = \frac{1}{n}\sum_{i=0}^{n} HHI(x_i) = 0.33$$

This scenario points to a fundamental relationship between the HHI computed for adverts, over the whole population, and the CAPI. Under the assumption that each person sees the same number of adverts, we find that:

$$HHI \leq CAPI \leq 1$$

In other words, if the HHI shows that a diverse set of adverts are being delivered, it is always possible that despite this, each person is effectively experiencing the same personal environment that they would in a monopoly and only receiving adverts from one seller in a particular market. However, if the HHI shows that there is an effective monopoly, this will always cause the CAPI to be high.

Moreover, the lower bound of $HHI = CAPI$ always occurs when the market is not separated into different subgroups, and everyone receives the same adverts at the same rate, and the upper bound of $CAPI = 1$ only occurs when each person receives adverts from only one advertiser.

Step 3: Working with the CAPI

To demonstrate the value of CAPI screening, we now consider a market with three mattress stores that are less equally distributed.[3] Imagine a three-stage scenario: First, there is only one incumbent ('Dreamy'). Dreamy advertises randomly to 50% of the population, not using targeted advertising. In stage 2, a competitor ('MyMattress') enters the market. Instead of advertising randomly, MyMattress uses some targeting to reach a specific half of the population, for example, women. In stage 3, another competitor ('SleepTight') appears. SleepTight deploys an extreme targeting strategy and only targets a specific tenth of the population. Let us assume they only target those consumers who they know to fare badly at calculating delayed costs. By placing salient adverts for consumer loans to purchase its mattresses, SleepTight thus engages in a form of 'behavioural manipulation' (see section 'Personalisation, consumer welfare and law enforcement').

We further assume that targeting positively affects the success of online adverts, increasing click-through rates. For Dreamy's non-targeted ads, we assume a click-through rate of 25% (i.e., every fourth consumer who is exposed to their advert will click on it). For MyMattress, we stipulate that their targeting leads to a click-through rate of 50% in the targeted population. For SleepTight, the click-through rate is 100% in the targeted population. If – and this will become important in the next section – MyMattress or SleepTight should happen to show their advert to a consumer outside of their target group, we assume their click-through rate to reverse to that of Dreamy (25%), i.e., their targeting is only effective towards the targeted consumers.

For these scenarios, we have calculated the following metrics, which are summarised in Table 1.

Each row describes one stage of the scenario, and the last row shows the percentual changes of all metrics moving from stage 2 to stage 3, i.e., the effect of the second competitor, SleepTight, entering the market.

Obviously, in a one-firm market, we have an HHI of 1 and a CAPI of 1 (see section 'Step 2: arriving at the CAPI'). Once the first competitor enters, the HHI halves. This is expected as the same market is now shared by two firms that effectively advertise to every second person in the population. The CAPI, however, does not halve but

Table 1. Changes in concentration become visible when examining the variance of the CAPI.

		Personal level	
Scenario	HHI	CAPI	Var(CAPI)
One firm	1.000	1.000	0.000
Two firms	0.500	0.829	0.056
Three firms	0.422	0.786	0.069
Change (2–3)	−16%	−5%	23%

Abbreviations: CAPI: Concentration-after-Personalisation Index; HHI: Herfindahl–Hirschman Index.

instead decreases only by roughly 16%. The gap between the HHI and the CAPI shows the latter's response to MyMattress' moderate targeting.

In the third row (the third stage of our scenario), we see a 16% decrease of the HHI due to the addition of a third advertiser, SleepTight. The CAPI only decreases by 5%, which already indicates that the advertisement for SleepTight is significantly differently distributed from the adverts of the two other advertisers, Dreamy and MyMattress. The increasing gap between the HHI and the CAPI signals a risk of consumer harm which is not detected by the competition-law metric of the HHI and a potential need for further regulatory scrutiny. This becomes even more apparent when examining the change in the variance of the CAPI, which has increased by 23%. We select the variance as a metric as it reflects the distribution of the CAPI, which is thus well-suited to detect outliers in the distribution. In this case, these outliers are the 'targeting pockets' we set out to detect in this paper, as they signal a high risk of harm to consumer welfare. The CAPI's variance shows that in the market with three mattress sellers, some people are experiencing a distinctly different exposure to targeted adverts than in the scenario with two firms.

As mentioned, a CAPI score is not evidence for a breach of law but provides summary statistical evidence that consumer welfare is at risk because of a lack of diversity in advertising exposure. Especially those 'targeting pockets' created by SleepTight can mean less consumer choice and an increased risk of price discrimination beyond the socially optimal as well as targeting with persuasive ads (see section 'Personalisation, consumer welfare and law enforcement'). Calculating the CAPI's variance thus helps to detect the aggressive targeting practices of SleepTight. Section 'Personalisation, consumer welfare and law enforcement' has argued how valuable such a discovery can be, not least because public oversight of advertising is compromised due to the fragmentation of consumers by targeted adverts.

Consequent consideration: Lifting the veil of group membership

Now that we detected a 'targeting pocket' with the CAPI, we consider a regulator's potential desire to know which consumer group sits in this pocket. The CAPI score alone, however, does not provide us with such information. To learn which consumers are trapped in the targeting pocket, the regulator (or auditor acting under novel regulation, see section 'Preliminary consideration: access to data') would need to obtain information about group distinction. Below, we assume that this information is available to the regulator. However, even if group information is unavailable, from a governance perspective, it is positive to note that the detection of the 'targeting pocket' has been successful without it. This is important as information about groups may be proprietary or, if consumers are fragmented by AI-driven technology, unintelligible to humans.

We expand Table 1, and the newly added columns show the CAPI for 10 groups of consumers, one of which is the group targeted by SleepTight. As in the previous section, we show the variance of this group-level CAPI (Table 2).

The values for the group level show that in this case, the change in the variance of the group-level CAPI matches the change in the variance of the personal-level CAPI. This is because, in our example, regulators perfectly identified the group that is targeted by SleepTight. The variance of the group-level CAPI can thus function as a benchmark as to whether the correct group has been identified. It further provides a benchmark for the policy solution we are going to suggest in the next section.

A novel solution: Calculating the optimal degree of added noise with the CAPI

In the previous section, we used the CAPI to detect the effects of targeted advertising on consumer choice that would remain undetected by the competition-law metric of the HHI, in particular, 'targeting pockets'. In this section, we take the identified gap between the CAPI and the HHI as a motivation to suggest a regulatory solution based on the CAPI: adding noise to targeting, i.e., exposing consumers to non-targeted, hence noisy, adverts.

We begin by arguing that as long as targeted advertising is not completely banned, 'targeting pockets' may still occur, even if the range of targeting variables is limited either by law or by an online advertiser's privacy policy (see section 'Regulatory options and second-generation tracking'). We then make the case for adding noise to targeting as a normative solution, at least if only (economic) consumer welfare is considered. As mentioned, other concerns such as privacy may suggest different solutions. We show that within this constraint, adding noise can be treated as an optimisation problem (see section 'Adding

Table 2. The variance of the group-level CAPI helps us identify which group has been targeted.

		Personal level		Group level	
Scenario	HHI	CAPI	Var(CAPI)	Group CAPI	Var(Group CAPI)
One firm	1.000	1.000	0.000	1.000	0.000
Two firms	0.500	0.829	0.056	0.776	0.056
Three firms	0.422	0.786	0.069	0.757	0.069
Change (2–3)	−16%	−5%	23%	−2%	24%

Abbreviations: CAPI: Concentration-after-Personalisation Index; HHI: Herfindahl–Hirschman Index.

noise as an optimisation problem'). As such, it responds well to the demands of the proportionality principle, a legal test that any regulatory mandate to add noise to targeting would have to pass. We then come back to our exemplary scenario of the online market for mattresses developed in the section 'A novel metric: the CAPI' and compute the optimal degree of added noise relative to the interests of consumers and sellers (see section 'Determining the optimal degree of noise with the CAPI').

Regulatory options and second-generation tracking

Currently, there are calls to ban certain uses of behavioural targeting in the USA and the EU by law (Jacobs, 2022; Kelly, 2022). As long as these bans will only limit the use of certain tracking variables such as those based on sensitive personal data (see Introduction), we do not expect this to eliminate all personalisation of advertising or other forms of personalised consumer interaction online, not even to speak of personalised pricing strategies.

At the same time, removing targeting categories by law and thus only allowing less fine-grained personalisation than would be technically possible is a less stringent alternative to a complete ban (Milano et al., 2021). It is thus normatively attractive insofar as it protects the economic interests of advertisers and seems to be the favoured option by EU policymakers. The European Commission's recently proposed regulation on political advertising intends to limit the use of sensitive information such as race, political opinions and religious beliefs for targeting purposes (European Commission, 2021b). The DSA is likewise bound to ban targeting based on sensitive personal data (European Parliament, 2022). In November 2021, Facebook announced to remove sensitive targeting options which reference, for example, public figures that relate to health, race or ethnicity, political affiliation, religion, or sexual orientation (Meta Platforms, Inc, 2021; for the legal problems which Facebook encountered, see Wachter, 2021).

While protecting consumers' privacy, this approach carries its own risk of being under-inclusive when considering economic harm to consumers. Selecting which categories are targetable and which are not requires normative choices. The law applicable to advertising in the EU – particularly EU consumer law – demonstrates this problem well, as it has neatly-defined characteristics of vulnerability in consumers. Article 5(3) UCPD lists 'mental or physical infirmity, age or credulity' as factors of vulnerability. This legal stipulation does not reflect the fact that all consumers can be vulnerable at some point in time (Trzaskowski, 2016). In as much as personalisation techniques can predict psychological states and cognitive biases, every consumer is the potential subject of economic exploitation in the market, for example by paying higher prices (Trzaskowski, 2016). There have thus been calls to reform the UCPD accordingly and widen the legal definition of consumer vulnerability (Hacker, 2021; Helberger et al., 2021; Laux et al., 2021a; Trzaskowski, 2016). However, as the inferential analytics behind personalisation is a fast-moving field of technology, there is always a residual risk of being under-inclusive when redefining legal terms.

Below, we show that for targeted advertising, there is an alternative regulatory solution that protects consumers' economic decision-making while being less restrictive and requiring less categorical normative choices than a legal ban: adding noise to targeting and thus randomly diversifying the exposure to adverts by an average individual consumer (on adding noise, see also Milano et al., 2021: 469–70).

Adding noise as an optimisation problem

To reiterate, we are endorsing noisy targeting because of the ambiguous welfare effects of targeted advertising. Targeting will hurt some consumers while others will benefit (see section 'Personalisation, consumer welfare and law enforcement'). A ban would, however, erase all potential welfare benefits that may result from an improved matching of adverts with consumer preferences and their price sensitivity. If consumers receive a certain amount of non-targeted, hence noisy, adverts, it allows them to see a wider range of offers (see also Laux et al., 2021a). They could thus become aware of the existence, quality and prices of products and services which advertisers would not pay a premium to supply them with based on their consumer profile. It would also allow them to see adverts that other consumer groups are being served with and thus increase public oversight.

Adding noise requires, however, a normative choice as to how much noise should be added to balance advertisers' and consumers' interests. This is essentially an optimisation problem, and as such, it responds well to the demands of the proportionality principle, which eventually every regulatory intervention through EU law would have to meet. Proportionality is one of the general principles of EU law and thus permeates the entire EU legal order (Tridimas, 2007). When applied, it often results in a structured test, where courts determine whether a (regulatory) policy pursues a legitimate aim, whether the measure is apt to achieve this aim, whether the measure is necessary to achieve this aim, and whether there is a proportionate balance struck between the aims pursued by the policy and its consequences, not least for the realisation of other legitimate interests (Young and De Búrca, 2017: 140).

We consider three interests that are affected by targeted advertising and are protected by the EU legal order, hence constituting 'legitimate aims' in a proportionality test First, exercising rational consumer choice (cf. sections 'Diversity in consumer choice', 'Persuasive advertising and rational consumer choice' and 'Price discrimination') is an interest protected by EU consumer law, which safeguards consumers against misleading advertising (cf. Art. 2(e) UCPD).

Second, as targeting fragments the audience of adverts, it can obstruct public oversight over adverts (Milano et al., 2021). The European public has a legally protected interest in having access to information (Court of Justice of the European Union, 2014: para 97). This should arguably also include some information about what commercial communication other European consumers are receiving, not least with a view to citizens' ability to file complaints about misleading advertising.

Third and finally, the advertiser has an economic interest to target consumers efficiently, save costs, and increase sales. The economic interests of sellers are, for example, recognised by Articles 15–17 of the EU Charter of Fundamental Rights. Moreover, the legitimacy of businesses' economic interests is broadly recognised in secondary EU law. For example, Article 15 of the General Data Protection Regulation (GDPR) can be limited for the sake of intellectual property rights, trade secrets and business interests. For simplicity, we are treating consumers' interest to receive advertising which matches their true preferences as coextensive with the advertisers' interest to target consumers.

Proportionality requires balancing these three interests against each other. The proportionality principle can be translated into an optimisation principle, hence treating legal interests as fulfillable to different degrees and requiring realisation to the highest degree that is actually and legally possible (Alexy, 2000, 2002: 44–110; for a critique, see Habermas, 1996: 258–59). We now show how the CAPI can be used to identify the optimal degree of added noise.

Determining the optimal degree of noise with the CAPI

Let us come back to our scenario with three mattress sellers. We are now introducing noise to their advertising as a regulatory response to the risks associated with the identified 'targeting pocket' (see section 'A novel metric: the CAPI').[4] We further balance the economic interest of firms, i.e., their click-through rates, with the consumers' and public's interest in diversity in advertising exposure.

The 'costs' of adding noise are understood as a loss of clicks compared to a scenario without noisy targeting. However, implementing noise in the current (yet contingent) market structures of the ad-tech industry would raise further questions about the distribution of 'costs' now understood as a loss in revenue. Several scenarios are imaginable. If noisy ads are shown on a publisher's website and generate fewer clicks than non-noisy targeted ads and the publisher receives more payment for ads which are clicked on, then the publisher will incur a large proportion of the costs of adding noise. One could further imagine that if publishers rely on ad-intermediaries to auction off their advertising space, they are now willing to pay less for the intermediaries' services who must implement noise by law, as their expected payoffs are now lower. However, the current market for ad-intermediary services is highly concentrated (see the references in Laux et al., 2021a: 731), making it unlikely that the dominant intermediary would lower prices on the publisher's side of the market. At the same time, if advertising budgets and publishers' space for adverts remain the same, noisy targeting could also increase prices in auctions as advertisers now need to place more ads to use up their budgets, thus keeping losses in revenue for publishers lower than in the scenario described before. That said, our interest lies with the consumer's perspective. The aim here is to show that the CAPI allows the balancing of click-through rates with diversity in advertising exposure.

Figure 1 represents both these metrics for varying levels of noise, from 0% to 100% added noise. Of course, 0% of noise describes the market as depicted in the section 'A novel metric: the CAPI'. 100% added noise prevents all sorts of targeting and makes now both Dreamy and MyMattress randomly advertise to 50% of the population (remember that MyMattress targeted a specific half of the population). SleepTight now advertises randomly to 10% of the population. As mentioned, outside of their target population, MyMattress' and SleepTight's click-through rates revert to that of the incumbent Dreamy (see section 'Step 3: working with the CAPI').

Now we observe that as noise increases, click-through rates deteriorate for the two competitors who deploy targeted advertising. For the incumbent Dreamy, however, there is no change visible – but for some irrelevant modelling fluctuation. Obviously, as SleepTight's previous click-through success was due to their targeting (every ad shown to their target group results in a click), with increasing noise more and more of their ads are shown outside of their target group where its click-through rate reverses to that of Dreamy. As long as SleepTight does not begin to show more ads and keeps on reaching a now-random 10% of the population, this trend continues.

Interestingly, the variance of the group-level CAPI (see section 'Consequent consideration: lifting the veil of group membership'), does not show a linear deterioration with increasing noise – the curve flattens out around the value of 40%. This suggests that there is an optimal degree of added noise, relative to our normative goals of protecting consumer choice and the economic interests of advertisers.

Selecting this noise level of 40% would indicate a 96% decrease in the variance of the group-level CAPI, no significant change in the click-through rates of Dreamy, a loss of click-through of 14% for MyMattress, and, respectively, 29% for SleepTight. The change of all other metrics along the different noise levels is listed in Table 3.

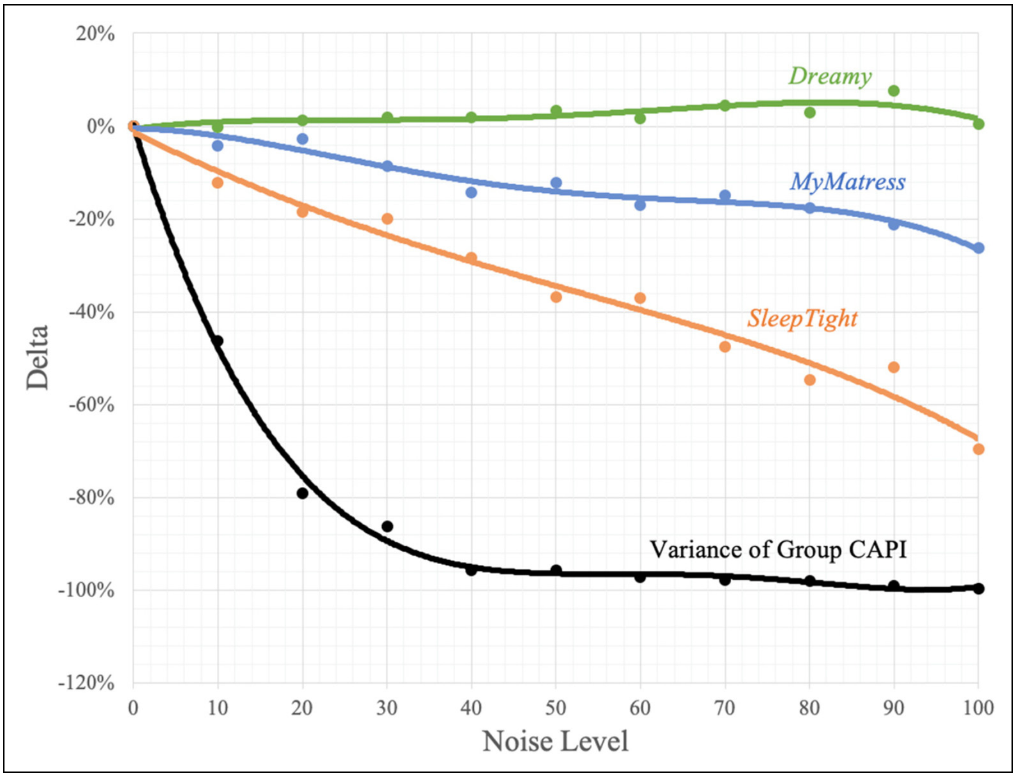

Figure 1. With increasing noise we see how the variance of the group-level CAPI diminishes quickly while click-through rates deteriorate significantly slower.

We can obviously only speculate about a court's assessment of the proportionality of adding 40% of noise. However, a 96% decrease in the reduction of the variance of the group-level CAPI at the 'price' of a 14% and 29% click-through loss appears well-balanced, given the risks associated with certain consumers being trapped in a 'targeting pocket'.

Moreover, as the adding of noise would have to be done by advertising technology providers, the lack of necessary group information (see section 'Step 3: working with the CAPI') on the side of regulators is curable: advertisers and ad-tech providers could simply be mandated to add the optimal level of noise, as determined by the CAPI. Yet even if mandated by law, the CAPI could not heal the asymmetry of access to data between the ad-tech industry and regulators, auditors and researchers.

Conclusion: Personalisation beyond advertising

In this paper, we introduced the CAPI as a novel metric with which to measure the effects of personalisation on consumer choice and consumer welfare. We used the example of targeted online advertising to show how the CAPI can measure the diversity of adverts experienced by an average individual. We showed how the CAPI screens the distribution of personalised adverts for concentration and detects 'targeting pockets' in which consumers can become trapped. Without such a screening tool, high degrees of concentration will remain hard to detect, not least because public oversight of advertising is compromised due to the fragmentation of consumers by targeting. We have argued elsewhere that the exposure to non-personalised options is important for consumer choice (Laux et al., 2021a); the European Parliament likewise appears to have recognised its importance when it decided that large online platforms should provide at least one recommender system that is not based on profiling (European Parliament, 2022). We further developed a novel policy solution to dissolve targeting pockets, i.e., the adding of noise to targeting. Based on the CAPI, we could show that there is an optimal degree of added noise, which successfully balances consumers' and advertisers' interests.

We anticipated the regulatory development around targeted advertising and demonstrated how the CAPI can help to strengthen the EU's coming DSA, its future political advertising regulation, its proposed AI Act, as well as existing EU consumer law. The CAPI provides a monitoring tool which helps to focus the auditing of advertising to areas of advert delivery in which the risk of consumer harm is heightened.

At this point, one may question whether research findings on 'filter bubbles' show that we are overestimating the likelihood and harmfulness of targeting pockets.

Table 3. Our model allows us to trade off a reduction in variance of the group-level CAPI with losses in click-through rates, finding an optimal level of noise around 40%.

	Deltas of Metrics					Deltas in Clicks		
Noise Level	*HHI*	*CAPI*	*Var(CAPI)*	*Group CAPI*	*Var(Group CAPI)*	*Dreamy*	*MyMatress*	*SleepTight*
0	0%	0%	0%	0%	0%	0%	0%	0%
10	1%	2%	-4%	-10%	-46%	0%	-4%	-12%
20	1%	0%	-2%	-22%	-79%	1%	-3%	-19%
30	0%	0%	-5%	-28%	-86%	2%	-9%	-20%
40	0%	3%	-9%	-35%	-96%	2%	-14%	-29%
50	1%	2%	-6%	-37%	-96%	3%	-12%	-37%
60	-1%	3%	-9%	-39%	-97%	2%	-17%	-37%
70	0%	-2%	-2%	-42%	-98%	5%	-15%	-48%
80	0%	1%	-5%	-42%	-98%	3%	-18%	-55%
90	-1%	0%	-4%	-45%	-99%	8%	-21%	-52%
100	2%	1%	-8%	-43%	-100%	0%	-26%	-70%

Note: Green is desirable, yellow is neutral, and red is undesirable.
Abbreviations: CAPI: Concentration-after-Personalisation Index; HHI: Herfindahl–Hirschman Index.

Actual click-through rates of targeted adverts are not nearly as high as assumed in the section 'A novel solution: calculating the optimal degree of added noise with the CAPI' and likely still remain below 1% (cf. Zuiderveen Borgesius et al., 2016: 10). The dangers of filter bubbles as a result of algorithmic personalisation have been found to be less harmful to people's access to information as originally thought – largely, however, because the technology is still insufficient and people are able to actively seek out other sources of information (Zuiderveen Borgesius et al., 2016: 10). We would like to offer two rejoinders. First, as a measure of concentration, the CAPI cannot appreciate the behaviour of consumers. It provides a screening tool which scans for areas with heightened risk in the distribution of personalised content, but not a conclusive answer as regards the occurrence of consumer harm. Moreover, targeting pockets are the extreme case based on which we developed our idea for the CAPI. This does not mean, however, that the CAPI is useless unless it detects a targeting pocket. Just like the HHI is a measure of market concentration and not only a detector of monopolies, the CAPI is a measure of concentration in personalisation, which may have negative effects on consumer welfare even below the level of targeting pockets.[5]

Second, personalisation technology may progress and permeate further domains of economic activity, actualising a need to revisit the findings on filter bubbles. After all, targeted advertising is one salient example of personalisation. Recommender systems and some forms of price discrimination are further examples. In the Introduction, we stated the risks of algorithmic collusion in setting prices, for which first empirical evidence is emerging. With the ongoing adoption of AI, consumers (and citizens) will likely increasingly interact with firms (and governments) through personalised communication and content. Behavioural manipulation through exploiting vulnerabilities in consumers may find its way directly into the composition of products and services (Acemoglu, 2021: 2). Screening tools such as the CAPI, which is not limited in its application to targeted adverts, should thus become important parts of regulators' toolboxes.

Lastly, metrics such as the CAPI demonstrate the importance of granting broader access to personalisation data for effective oversight and auditing (Laux et al., 2021b). The CAPI could easily be utilised by different actors, such as regulators, auditors and industry itself. Such broad utilisation would be beneficial even where access to data remains unequally distributed. As shown above, if information about targeted groups remains inaccessible to auditors, regulators could still mandate industry to add optimal degrees of noise as determined by the CAPI.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This article is a deliverable of the 'A right to reasonable inferences in advertising and financial services' project. This project is financially supported by a research grant from the Miami Foundation, programmatic funding from Luminate Group, and research grants from the British Academy (grants PF\170151 and PF2\180114). This work has also been supported through research funding provided by the Wellcome Trust (grant nr 223765/Z/21/Z), Sloan Foundation (grant nr G-2021-16779), and the Department of Health and Social Care (via the AI Lab at NHSx).

ORCID iDs

Johann Laux https://orcid.org/0000-0003-3043-075X
Fabian Stephany https://orcid.org/0000-0002-0713-6010

Chris Russell https://orcid.org/0000-0003-1665-1759
Sandra Wachter https://orcid.org/0000-0003-3800-0113
Brent Mittelstadt https://orcid.org/0000-0002-4709-6404

Notes

1. Critics of the informative view hold that in the digital age, consumers do not need advertising anymore to collect information about products they are interested in (Woodcock, 2018: 2274–2275). What remains is mainly persuasive advertising, which constitutes "socially wasteful spending" (Woodcock, 2018: 2337). Moreover, a third view on why consumers respond to advertising holds adverts to be complementary to the advertised product. For example, those consumers who value 'social prestige' will find a product that is appropriately advertised as generating greater prestige in its consumption (Bagwell, 2007: 1706).
2. The HHI has been adapted before by political scientists to measure the effective number of parties in a political system in the so-called L-T index, which accounts for the fact that not all parties are of the same size (following Laakso and Taagepera, 1979; see further Grofman, 2019: 149, fn. 7). The L-T index allows to describe the fragmentation of the party system according to the electoral weight of the parties in elections (Dodeigne and Pilet, 2021: 237).
3. We created a model with a population of 1000 potential consumers, which we segmented into 10 equally sized groups (A-J). Exposure to advertisement is labelled with a value of 1 and 0 otherwise. Hence, for Dreamy, we distributed values of 0 and 1 randomly across all 1000 individuals. For MyMattress, we assigned all 500 individuals in groups A–E a value of 1, and 0 for the individuals in groups H–J. For SleepTight, we assigned all 100 individuals in group A a value of 1, and 0 for the remaining 900 individuals. We did not distribute adverts through auctioning, the current practice in online advertising; cf. the references in Laux et al., 2021a: 731. With auctioning, we would expect the chance for 'targeting pockets' to occur to be at least the same, if not even higher, as it is in our model.
4. With the introduction of different levels of noise, we gradually shifted away from the original advertisement distributions in our model (cf. n 1) and substituted them with random noise. However, we kept the initial size of the advertised population for each firm intact. This means that, throughout all degrees of noise, Dreamy and MyMattress advertise to 500 individuals, while SleepTight advertises to 100 individuals.
5. This may vary from market to market and will further depend on consumer behaviour. For example, how likely are consumers to seek out information on insurance besides the offers they are being advertised with?

References

Acemoglu D (2021) Harms of AI. NBER Working Paper 29247. Available at: https://economics.mit.edu/files/21848?utm_campaign=Events%3A%20Economic%20Studies&utm_medium=email&utm_content=188861446&utm_source=hs_email (all websites accessed on 5 April 2022).

Aguirre E, Roggeveen AL, Grewal D, et al. (2016) The personalization-privacy paradox: Implications for new media. *Journal of Consumer Marketing* 33(2): 98–110.

Alexy R (2000) On the structure of legal principles. *Ratio Juris* 13(3): 294–304.

Alexy R (2002) *A Theory of Constitutional Rights*. Oxford: OUP.

André Q, Carmon Z, Wertenbroch K, et al. (2018) Consumer choice and autonomy in the age of artificial intelligence and big data. *Customer Needs and Solutions* 5(1): 28–37.

Assad S, Clark R, Ershov D, et al. (2020) *Algorithmic pricing and competition: Empirical evidence from the German retail gasoline market*. CESifo Working Paper No. 8521, CESifo, Munich.

Australian Consumer and Competition Commission (ACCC) (2019, 26 July) *Digital platforms inquiry*. Final Report June 2019, Australian Consumer and Competition Commission, Canberra.

Bagwell K (2007) The economic analysis of advertising. In: Armstrong M and Porter R (eds) *Handbook of Industrial Organization*, Vol. 3. Amsterdam: Elsevier, pp. 1701–1844.

Balebako R, Leon PG, Shay R, et al. (2012) Measuring the effectiveness of privacy tools for limiting behavioral advertising. In: Web 2.0 Security and Privacy 2012, San Francisco, CA, USA, 24 May 2012. New York: Institute of Electrical and Electronics Engineers.

Bar-Gill O (2019) Algorithmic price discrimination. *University of Chicago Law Review* 86(2): 217–254.

Bourreau M, De Streel A and Graef I (2017) Big data and competition policy. Report, Centre on Regulation in Europe, Brussels, 16 February.

Calo R (2014) Digital market manipulation. *George Washington Law Review* 82(4): 995–1051.

Calvano E, Calzolari G, Denicolò V, et al. (2020) Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review* 110(10): 3267–3297.

Court of Justice of the European Union (2014) Case C-131/12 *Google Spain SL and Google Inc. v. Agencia Española de Protección de Datos (AEPD) and Mario Costeja González*, published in the electronic ERC.

Cozza V, Hoang VT, Petrocchi M, et al. (2016) Experimental measures in news personalization in Google News. In: Casteleyn S, Dolog P and Pautasso C (eds) *Current Trends in Web Engineering*, Vol. 9881. Cham: Springer, pp. 93–104.

Cyphers B (2021) Google's FLoC is a terrible idea. Available at: https://www.eff.org/deeplinks/2021/03/googles-floc-terrible-idea (accessed 26 August 2022).

Dodeigne J and Pilet JB (2021) Centralized or decentralized personalization? Measuring intra-party competition in open and flexible list PR systems. *Party Politics* 27(2): 234–245.

Dos Santos JC, Siqueira SWM, Pereira N, et al. (2020) Is there personalization in Twitter search? A study on polarized opinions about the Brazilian welfare reform. In: 12th ACM Conference on Web Science, 6–10 July 2020, pp. 267–276. New York: Association for Computing Machinery.

European Commission (2009) Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings. OJ C 45/02. Brussels, 24 February.

European Commission (2020a) On the European democracy action plan. Communication from the Commission to the European Parliament, the Council, the European Economic

and Social Committee of the Regions. COM(2020) 790 final. Brussels, 3 December.
European Commission (2020b) Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC. COM(2020) 825 final. Brussels, 15 December.
European Commission (2021a) Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. COM(2021) 206 final. Brussels, 21 April.
European Commission (2021b) European Democracy: Commission sets out new laws on political advertising, electoral rights and party funding. Available at: https://ec.europa.eu/commission/presscorner/detail/en/ip_21_6118.
European Commission (2021c) Proposal for a regulation of the European Parliament and of the Council on the transparency and targeting of political advertising. COM(2021) 731 final. Brussels, 25 November.
European Parliament (2022) Digital services act: Regulating platforms for a safer online space for users. Available at: https://www.europarl.europa.eu/news/en/press-room/20220114IPR21017/digital-services-act-regulating-platforms-for-a-safer-online-space-for-users.
Ezrachi A (2018) EU Competition law goals and the digital economy. Oxford Legal Studies Research Paper No. 17/2018. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3191766.
Ezrachi A and Stucke ME (2016a) Virtual competition. *Journal of European Competition Law & Practice* 7(9): 585–586.
Ezrachi A and Stucke ME (2016b) *Virtual Competition*. Cambridge, MA: Harvard University Press.
Filippino M, Masters B, Hodgson C, et al. (2021) Apple's privacy policy wreaks havoc on rivals. Financial Times, 2 November. Available at: https://www.ft.com/content/d5d05648-ea0e-4c8a-ac0f-daad0d8ce5fb.
Goldfarb A (2014) What is different about online advertising? *Review of Industrial Organization* 44(2): 115–129.
Gonzales W (2022) How machine learning is shaping the future of advertising. Forbes, 18 January. Available at: https://www.forbes.com/sites/forbesbusinesscouncil/2022/01/18/how-machine-learning-is-shaping-the-future-of-advertising/?sh=285cec21361e.
Grant N (2021) Google to let android users opt out of tracking, following Apple. Fortune, 4 April. Available at: https://fortune.com/2021/06/03/google-android-users-opt-out-of-tracking-apple/.
Grofman B (2019) Political choices in one dimension: Theory. In: Congleton RD, Grofman B and Voigt S (eds) *The Oxford Handbook of Public Choice*, Vol. 1. Oxford: OUP, pp. 147–166.
Guha S, Cheng B and Francis P (2010) Challenges in measuring online advertising systems. In: 10th ACM SIGCOMM conference on Internet measurement, Melbourne, AUS, 1–30 November 2010, pp. 81–87. New York: Association for Computing Machinery.
Habermas J (1996) *Between Facts and Norms*. Oxford: Polity Press.
Hacker P (2021) Manipulation by algorithms: Exploring the triangle of unfair commercial practice, data protection, and privacy law. *European Law Journal*. Epub ahead of print 9 December 2021. DOI: 10.1111/eulj.12389.
Ham CD (2017) Exploring how consumers cope with online behavioral advertising. *International Journal of Advertising* 36(4): 632–658.
Hannák A, Sapieżyński P, Khaki AM, et al. (2017) Measuring personalization of web search. In: 22nd International Conference on World Wide Web, Rio de Janeiro, Brazil, 13–17 May 2017, pp. 527–538. Rio de Janeiro: International World Wide Web Conferences Steering Committee.
Hansen KT, Misra K and Pai MM (2021) Algorithmic collusion: Supra-competitive prices via independent algorithms. *Marketing Science* 40(1): 1–12.
Helberger N, Sax M, Strycharz J, et al. (2021) Choice architectures in the digital economy: Towards a new understanding of digital vulnerability. *Journal of Consumer Policy*. Epub ahead of print 22 December 2021. DOI: 10.1007/s10603-021-09500-5.
Jacobs H (2022) Is Momentum shifting toward a ban on behavioural advertising? The Markup, 3 February. Available at: https://themarkup.org/ask-the-markup/2022/02/03/is-momentum-shifting-toward-a-ban-on-behavioral-advertising.
Kelly M (2022) Democrats unveil bill to ban online 'surveillance advertising'. The Verge, 18 January. Available at: https://www.theverge.com/2022/1/18/22889903/democrats-targeted-advertising-facebook-google-surveillance.
Krafft TD, Gamer M and Zweig KA (2019) What did you see? A study to measure personalization in Google's search engine. *EPJ Data Science* 8(1): 38.
Laakso M and Taagepera R (1979) 'Effective' number of parties: A measure with application to west Europe. *Comparative Political Studies* 12(1): 3–27.
Laux J, Wachter S and Mittelstadt B (2021a) Neutralising online behavioural advertising: Algorithmic targeting with market power as an unfair commercial practice. *Common Market Law Review* 58(3): 719–750.
Laux J, Wachter S and Mittelstadt B (2021b) Taming the few: Platform regulation, independent audits, and the risks of capture created by the DMA and DSA. *Computer Law & Security Review* 43: 105613.
Le H, Maragh R, Ekdale B, et al. (2019) Measuring political personalization of google news search. In: The World Wide Web Conference 2019, San Francisco, CA, USA, 13–17 May 2019, pp. 2957–2963. New York: Association for Computing Machinery.
Marotta V, Wu Y, Zhang K, et al. (2021) The welfare impact of targeted advertising technologies. *Information Systems Research* 33(1): 131–151.
Matz SC and Netzer O (2017) Using big data as a window into consumers' psychology. *Current Opinion in Behavioural Sciences* 18(1): 7–12.
McGee P (2021) Snap, Facebook, Twitter and YouTube lose nearly $10bn after iPhone privacy changes. Financial Times, 31 October. Available at: https://www.ft.com/content/4c19e387-ee1a-41d8-8dd2-bc6c302ee58e.
Meaker M (2022) The EU has a plan to fix internet privacy: Be more like apple. Wired, 25 January. Available at: https://www.wired.com/story/eu-digital-services-act-apple/.
Meta Platforms, Inc. (2021) Removing certain Ad targeting options and expanding our Ad controls. Available at: https://www.facebook.com/business/news/removing-certain-ad-targeting-options-and-expanding-our-ad-controls.

Milano S, Mittelstadt B, Wachter S, et al. (2021) Epistemic fragmentation poses a threat to the governance of online targeting. *Nature Machine Intelligence* 3(6): 466–472.

Milmo D (2022) Google is changing how it tracks us online. The Guardian, 2 February. Available at: https://www.theguardian.com/technology/2022/feb/02/techscape-google-chrome-cookies.

Poort J and Zuiderveen Borgesius F (2021) Personalised pricing: The demise of the fixed price. In: Kohl U and Eisler J (eds) *Data-Driven Personalisation in Markets, Politics and Law*. Cambridge: Cambridge University Press, pp. 174–189.

Puschmann C (2018) Beyond the bubble: Assessing the diversity of political search results. *Digital Journalism* 7(6): 824–843.

Salehi S, Du JT and Ashman H (2015) Examining personalization in academic web search. In: 26th ACM Conference on Hypertext & Social Media 2015, Guzelyurt, Northern Cyprus, 1–4 September 2015, pp. 103–111. New York: Association for Computing Machinery.

Strahilevitz LJ (2013) Toward a positive theory of privacy law. *Harvard Law Review* 126(7): 2010–2042.

Tridimas T (2007) *The general principles of EU law*. 2nd ed. Oxford: OUP.

Trzaskowski J (2016) Lawful distortion of consumers' economic behaviour. *European Business Law Review* 27(1): 25–49.

United States Department of Justice (2018) Herfindahl-Hirschmann index. Available at: https://www.justice.gov/atr/herfindahl-hirschman-index.

Wachter S (2021) Affinity profiling and discrimination by association in online behavioural advertising. *Berkeley Technology Law Journal* 35(2): 367–430.

Weatherill S (2016) Consumer protection. In: Patterson D and Södersten A (eds) *A Companion to European Union Law and International Law*. Chichester: Wiley, pp. 285–295.

White LJ (2013) Market power: How does it arise? How is it measured? In: Thomas CR and Shughart WFII (eds) *The Oxford Handbook of Managerial Economics*. Oxford: OUP, pp. 31–65.

Woodcock RA (2018) The obsolescence of advertising in the information age. *Yale Law Journal* 127(8): 2270–2341.

Young AL and De Búrca G (2017) Proportionality. In: Vogenauer S and Weatherill S (eds) *General Principles of Law: European and Comparative Perspectives*. London: Hart, pp. 133–144.

Zuiderveen Borgesius FJ (2015) *Improving Privacy Protection in the Area of Behavioural Targeting*. Alphen aan den Rijn: Wolters Kluwer.