Hypertargeting, Limited Attention, and Privacy: Implications for Marketing and Campaigning*

Florian Hoffmann† Roman Inderst‡ Marco Ottaviani§

June 2013

Abstract

Using data collected on the internet, firms and political campaigners can tailor their communication to the preferences and orientations of individual consumers and voters, a practice known as hypertargeting. This paper models hypertargeting as selective disclosure of information to an audience with limited attention. We characterize the private incentives and welfare impact of hypertargeting depending on the wariness of the audience, the intensity of competition, and the feasibility of price discrimination. We show that stricter privacy regulation requiring user consent often backfires, but is unambiguously beneficial when consumers are naive, competition is limited, and firms are able to price discriminate.

Keywords: Hypertargeting, selective disclosure, limited attention, consumer privacy regulation, personalized pricing, competition.

JEL Classification: D83 (Search; Learning; Information and Knowledge; Communication; Belief), M31 (Marketing), M38 (Government Policy and Regulation).

*Inderst and Ottaviani acknowledge financial support from the European Research Council through ERC Advanced Grant 229921 (Inderst) and ERC Advanced Grant 295835 (Ottaviani).

†Johann Wolfgang Goethe University Frankfurt. E-mail: fhoffmann@finance.uni-frankfurt.de.

‡Johann Wolfgang Goethe University Frankfurt and Imperial College London. E-mail: inderst@finance.uni-frankfurt.de.

§Bocconi University, Milan. E-mail: marco.ottaviani@unibocconi.it.

"In the old days, everyone—Democrats, Republicans, enthusiasts, nonvoters and undecideds—saw the same television ads. Now the campaigns use 'big data' to craft highly customized and even personalized messages as people go from website to website. The campaigns test just the right ads for each voter. . . . A wealthy urban liberal sees different ads online than a working-class centrist. People who care more about jobs see different ads than people who focus on social issues." *L. Gordon Crovitz, How Campaigns Hypertarget Voters Online, Wall Street Journal, November 4, 2012.*

"A glimpse into the future of advertising: Jim, the Chief Marketing Officer of a consumer products company . . . never knew exactly who he was reaching or how effective his advertising was. . . . Previously, Jim bought broad-reaching spots, hoping to reach his target audience. But now, targeting, measurement and analysis capabilities that previously were only available for Web advertising are available for all channels. Jim can develop an interactive, integrated marketing plan tailored to his individual target consumer, and he pays based on actual impact rather than by cost per thousand impressions (CPM). His marketing message follows those customers across content platforms to deliver a consistent experience. His advertising includes a mix of creative spots and formats, like special interest content, product placement and self-published advertising that are tailored to his consumers' preferences, community affiliations and devices. This enables his target consumers—be they traditional moms in Des Moines, Iowa, urban professionals in Berlin or university students in South Korea—to better experience the value of his product. . . . Jim creates multiple versions of his advertising campaigns in order to appeal to numerous customer micro-segments. *IBM Global Business Services, The End of Advertising as We Know It, 2007.*

1 Introduction

Firms and political candidates have traditionally had two distinct ways to convey information and persuade consumers and voters. They could either broadcast their messages through old media (leaflets, billboards, newspapers, and television), thereby achieving only a coarse segmentation of the audience, mostly along channel types and regional boundaries. Alternatively, they could customize their communication strategy to single individuals or small groups with direct marketing and ground-game campaigning. Firms and candidates could gather critical knowledge about their audience and send tailored messages through face-to-face contact by an experienced salesperson or canvassing by a skilled campaigner.

Nowadays, developments in computer technology increasingly allow sellers and campaigners to systematically collect personal and detailed data about an individual's past purchasing behavior, browsing activity, credit history, as well as personal likes and dislikes shared on

social networking sites.[1] When conducting what might appear to be an impersonal transaction through the internet, a great deal of personal information can be used to finely target consumers and voters, a practice known as behavioral targeting or hypertargeting.[2]

The greater availability of personally identifiable data on the internet blurs or even reverses the traditional distinction between personal selling or campaigning and remote communication. However, concerns are often raised that some consumers and voters remain blithely unaware of this practice and might suffer as a result. An active debate is underway among policymakers about reforming the regulatory framework for consumer privacy with an emphasis on the collection and use of personal data on the internet. While in this area the U.S. currently relies mostly on industry self regulation, policymakers and Congress are considering stricter regulation of consumer privacy.[3] In recent years, European legislators have intervened more directly by raising barriers to the collection and use of personally identifiable data about past purchases or recent browsing behavior, including a requirement that firms seek explicit consent to collect information using so-called cookies.[4]

What is the welfare impact of mandating consumer consent to track online behavior? Motivated by this question, this paper models hypertargeting as the collection of information about consumer preferences (or voter orientations) that allows firms (or candidates) to tailor their communication by selectively disclosing attributes about their offerings. A key feature of our model is the presumption that the scope of firms' communication to consumers is naturally restricted by factors such as airtime and screen space, or simply by the limited

[1] Information can be either collected directly or acquired from search engines and specialized data vendors. In its privacy policy, Facebook writes: "We allow advertisers to choose the characteristics of users who will see their advertisements and we may use any of the non-personally identifiable attributes we have collected (including information you may have decided not to show to other users, such as your birth year or other sensitive personal information or preferences) to select the appropriate audience for those advertisements." *https://www.facebook.com/note.php?note_id=+322194465300*

[2] "Tailor your ads and bids to specific interests: Suppose you sell cars and want to reach people on auto websites. You believe that the brand of cars you sell appeals to a wide variety of people, but some of them may react more positively than others to certain types of ads. For example, . . . you could show an image ad that associates a family-oriented lifestyle with your car brand to auto website visitors who're also interested in parenting." *Google AdWords, http://support.google.com/adwords/answer/2497941?hl=en*

[3] See American Association of Advertising Agencies (2009) for a widely adopted set of self-regulatory principles for online behavioral advertising. On the U.S. policy debate, see White House (2012), Federal Trade Commission (2012), and the extensive discussion of the Do Not Track legislation proposals on wikipedia.

[4] See the Data Protection Directive (1995/46/EC) and the Privacy and Electronic Communications Directive (2002/58/EC), also known as the E-Privacy Directive, which regulates cookies and other similar devices through its amendments such as Directive 2009/136/EC, the so-called EU Cookie Directive, and the Privacy and Electronic Communications (EC Directive) (Amendment) Regulations 2011. The current prescription is that "cookies or similar devices must not be used unless the subscriber or user of the relevant terminal equipment: (a) is provided with clear and comprehensive information about the purposes of the storage of, or access to, that information; and (b) has given his or her consent." More recently, European authorities have been pressuring internet giants such as Facebook and Google to limit the collection of personal data without user consent.

attention of consumers. Given this limited attention, firms disclose different attributes of a product to different consumers so as to strategically affect their willingness to pay.

Provided that firms or their agents are able to gather the necessary information, they hypertarget consumers by selectively disclosing marketing information, say through advertising or a salesperson, about different attributes of a product or service depending on their perceived individual preferences. For example, depending on a consumer's recorded past purchases or recently visited sites, a firm could learn how much a consumer values style relative to comfort. A tailored message may then devote more space or airtime to displaying the stylish features of a product. Alternatively, the message may provide a more sombre check list of a product's user-friendly features.

Our positive and normative results depend on how rational consumers are. While wary consumers are able to make a partial inference about the feature that is not disclosed with hypertargeting, unwary consumers cannot make this inference. Selective disclosure induces an upward shift in the preferences of unwary consumers, but makes the preferences of wary consumers more dispersed.

A key insight we obtain is that selective disclosure may actually benefit consumers. In fact, when there are sufficiently many firms competing in the market, we show how selective disclosure can benefit consumers, regardless of whether they are wary or unwary of this practice and even when firms also engage in personalized pricing. Then, the overriding effect is that selective communication proves to be more informative in the following way. Intuitively, under selective disclosure, consumers not only learn about the attribute that is communicated, but also learn indirectly and partly about the attribute that is not communicated. Because firms selectively disclose the most favorable attribute, the undisclosed attribute must be less favorable.

When firms cannot personalize prices (for example, because prices are fixed due to consumer arbitrage or in the application to political campaigning), we show that wary consumers profit from selective disclosure. Competition ensures that consumers benefit also when the use of personally identifiable information allows firms to price discriminate. But competition benefits unwary consumers particularly through the following mechanism. Given that unwary consumers have inflated perceptions about the products offered by all firms practicing selective disclosure, the distortion in the processing of information obtained from one firm is compensated by a similar distortion of information obtained from competing firms. Through this channel, competition eliminates bias and thus protects unwary consumers.

The introduction of personalized pricing changes the outcome in important ways. In fact, the extent to which the efficiency gains associated with more informative communication are shared between firms and consumers depends on whether firms can price discriminate accord-

ing to the perceived expected valuation of a particular consumer. Such price discrimination may only be feasible for services or low-value products, when customers or intermediaries have little scope for arbitrage. When a firm is in a monopolistic position, price discrimination can result in exploitative behavior, making regulatory intervention desirable. With the introduction of competition, perceived product differentiation matters. Selective disclosure dampens competition by increasing perceived differentiation, from an ex-ante perspective. Consumer ignorance about selective disclosure reduces differentiation, spurs rivalry among firms for a particular consumer, and lowers prices. With competition and price discrimination, consumer unwariness becomes a blessing.

Taken together, our analysis suggests that hypertargeting—the collection and use of personally identifiable data by firms to tailor selective disclosure—should benefit consumers when they are adequately protected by at least one of the following three conditions: their own wariness, competition, or the inability of firms to practice personalized pricing. A strong rationale for regulation emerges only when all three conditions are not met, that is, when a monopolist practices both selective communication and personalized pricing to exploit unwary consumers. Otherwise, even seemingly light-touch regulation, such as requiring consumer consent to collect and use personal data, may backfire by giving firms a way to commit to avoid selective communication with wary consumers, who are made worse off as a result.

In an extension, we apply the model to political campaigning followed by voting for one of two candidates or options in a referendum.[5] While personal purchasing decisions depend on the preferences of each individual consumer in isolation, collective voting decisions depend on the aggregation of individual preferences. This key difference allows us to obtain a number of new insights. From the perspective of an individual voter, for instance, what matters now is the product of the probability with which the voter becomes pivotal and the conditional utility in this event. While the conditional utility always increases when more candidates communicate selectively, at least when voters are wary, selective disclosure by some but not all candidates can tilt the vote shares of different candidates, thus reducing the probability that any given voter becomes pivotal. The model also applies when there are differences of preferences or political orientation across groups of voters, so that selective communication targets different groups instead of each voter individually.

Our key departure from much of the literature on strategic disclosure, as initiated by Grossman (1981) and Milgrom (1981) and Milgrom and Roberts (1986), is that firms cannot communicate all the attributes they know. This may be due to space or time constraints

[5]Kendall, Nannicini, and Trebbi (2013) find causal evidence of the effect of campaign information on electoral outcomes in a large-scale field experiment. For a view from the advertising industry, see Abse's (2013) account of hypertargeting strategies pursued in the 2012 US presidential election.

or simply because (too much) information "consumes the attention of its recipients" (Simon 1971).[6] In our setting with limited attention, non-disclosure of an attribute does not trigger a complete unraveling of consumer perceptions, as is the case in a world with unlimited attention. Fishman and Hagerty (1990) and Glazer and Rubinstein (2004) also consider models with constraints on the disclosure of verifiable information, though their focus is different. In Glazer and Rubinstein (2004) it is the receiver who can choose which information is revealed. In contrast to the disclosure literature, including Fishman and Hagerty (1990), ours is essentially a model of horizontal differentiation. Given that consumers differ in their preferences, it is optimal for firms to communicate different attributes to different consumers.[7]

Because attention is constrained, a sender endogenously faces an opportunity cost from disclosing one attribute rather than another; thus, our baseline equilibrium construction relates to Jovanovic's (1982) model where disclosure is exogenously costly. In a disclosure setting in which the fraction of receivers who fail to update their beliefs following the lack of disclosure (analytical failure) is higher than the fraction of receivers who do not attend to the disclosure (cue neglect), Hirshleifer, Lim, and Teoh (2004) obtain an equilibrium in which the sender only discloses high realizations,[8] as in Jovanovic (1982). Our unwary consumers, instead, attend to the disclosed attribute but fail to make the appropriate inference about the undisclosed attribute, which is chosen selectively by the sender. Thus, relative comparisons across different dimensions of information play a key role in our model.[9]

Our distinction between hypertargeting with unwary and wary consumers broadly corresponds to the difference between persuasive and informative advertising.[10] In either case, in our model the choice of information policy is naturally constrained between being either selective or non-selective disclosure. In contrast, Rayo and Segal (2010) and Kamenica and Gentzkow (2011) analyze a sender's unconstrained choice of information policy to op-

[6]The limited capacity of individuals to process information is currently being investigated in a number of other areas, ranging from macroeconomics (e.g., Sims 2003) to organization economics (Dessein, Galeotti, and Santos 2012). In our model, it is the sender who must choose a particular attribute to disclose given the limitation of the communication channel, rather than the receivers having to choose how to optimally direct their limited attention and information processing capacity.

[7]Our analysis abstracts away from externalities across the communication strategies of firms due to congestion effects and information overload; see, for example, Van Zandt (2004) and Anderson and de Palma (2012) for analyses in this direction in models à la Butters (1977). See also Johnson (2013) for a welfare analysis of the impact of targeted advertising in the presence of advertising avoidance by consumers.

[8]This is, more generally, akin to the behavior of agents in the "cursed equilibrium" concept of Eyster and Rabin (2005).

[9]Relative comparisons across dimensions also play a role in the construction of cheap-talk equilibria by Chakraborty and Harbaugh (2007, 2010) and Che, Dessein, and Kartik (2013).

[10]See, for example, Dixit and Norman (1978) and Grossman and Shapiro (1984) on the welfare impact of persuasive and informative advertising, and Bagwell (2007) for a systematic survey on the economics of advertising.

timally tweak the beliefs of a rational receiver.[11] Section 4.3's setting with voting relates to the problem of persuading a group to take a collective decision considered by Caillaud and Tirole (2007); however, in our model voters cast their ballot simultaneously rather than sequentially.[12]

Also focusing on horizontal differentiation, and thus on information about individual suitability, Lewis and Sappington (1994) consider a seller's incentives to provide consumers with better information, restricting attention to a "truth-or-noise" information structure.[13] Johnson and Myatt (2006) define information quality in terms of a rotation of the posterior distribution of consumers' expected valuations. We identify conditions for when the use of selective communication induces such a rotation, so that the distributions with and without selective disclosure satisfy a single-crossing condition. The rest of our analysis is then based on the use of single crossing, which allows us to obtain clear-cut results on how selective disclosure affects consumer surplus. Our analysis of the model with personalized pricing is also related to Ganuza and Penalva's (2010) work on the incentives for information provision in a second-price auction, as explained in Section 5.

The paper proceeds as follows. Section 2 introduces our baseline model. Section 3 develops the key comparison between the distributions of consumer valuations induced by selective and non-selective communication. Section 4 analyzes markets in which there is no scope for personalized pricing and develops an application to political campaigning. Section 5 turns to markets for which personalized pricing is possible. Section 6 concludes by reviewing our results in the light of a more detailed discussion of related literature in economics and marketing. Appendix A collects some omitted proofs. Supplementary Appendices B and C report additional material mentioned in the main text.

2 Baseline Model

At the heart of our analysis is a game of communication between senders and receivers. For concreteness, we frame the baseline model in the context of firms interacting with consumers; in Section 4.3 we enrich the model to analyze the case of candidates campaigning for voters. Firms must communicate truthfully, for example, because the transmitted information is verifiable and mendacious statements result in prohibitive losses of reputation or liability.

[11] See also DellaVigna and Gentzkow (2010) for a survey of the literature on persuasion across economics, marketing, and political science.

[12] Our model abstracts away from direct costs of campaign advertising. A political economy literature has focused on the welfare economics of contribution limits for financing campaign advertising; see, for example, Coate (2004) and Prat's (2007) survey.

[13] See also Kamenica, Mullainathan, and Thaler (2011) for a discussion of situations in which firms might know more about consumer preferences than consumers themeselves do.

In line with the literature, we refer to this communication as disclosure. Disclosure affects a consumer's perceived valuation of a firm's product relative to an alternative option. Two cases will be of interest. In the case of monopoly, the alternative for the consumer is an outside option of known value. With competition, the alternative is to purchase from a different firm. In both cases, the value of the outside option is unaffected by a firm's disclosure strategy.

Each firm may choose two different disclosure strategies. A firm discloses product attributes either selectively or non-selectively to each individual consumer. This section sets the stage by focusing on how individual consumers update their beliefs (and thus their demand) when faced with selective or non-selective disclosure by a single firm. Building on the simple analysis of the effect of selective disclosure on consumer demand, in the rest of the paper we endogenize the firm's choice of information acquisition and disclosure strategy, while also embedding the model into a market environment that allows for multiple firms and personalized pricing.

Persuasion through selective disclosure is relevant in a number of settings in which the firm or its agents learn about the preferences of consumers:

- First, the model applies to traditional marketing communication strategies. As we mention in the opening paragraph of the introduction, old media allow for a segmentation of consumers into coarse groups; nevertheless, different messages can be (and often are) sent to groups with different preferences.[14]

- Second, consider a face-to-face interaction between a salesperson and an individual consumer. Even when meeting a consumer for the first time, an experienced salesperson should be able to draw inferences about the consumer's needs and preferences and use the limited time available (or the consumer's limited attention) to communicate only those product attributes that dovetail nicely with those preferences.

- A third relevant setting is distance selling through communication channels that were previously anonymous, but now allow for increased personalization given the ability

[14] An example in point is Wrangler jeans advertising strategy: ". . . while their main product is largely the same in feel, quality and color (commercial denim is pretty much commercial denim), the experience they market that product reflecting or emulating is very different across the globe. . . . In the US, Wrangler is perceived as the cowboy brand, the go-to jeans for the workers of the Midwest. These people aren't fashionable—and they frankly don't care. . . . This is why their tagline is 'Real. Comfortable. Jeans.' And Wrangler gives them that experience by associating their denim with their lifestyle through advertising. If you look at the European and Asian sites, you'll see a very different experience. While still 'American,' the brand is giving people in those countries a taste of the limitless American freedom of an open road and exploration. . . . You'll also notice that the tagline is 'Worn Across America'—as if the jeans represent the movers and shakers and nomads on the go and exploring the last frontier of the American West." *Listen Here, Sweetheart. It's ALL about Marketing, March 31, 2011, http://philliphess.tumblr.com/*

of firms to collect personally identifiable data on the internet. Based on an individual consumer's profile, a firm may choose how to best use the limited amount of time or space to selectively convey the attributes of a product. In what follows, we frame the discussion in terms of this third application. Our policy implications are then directed to the collection of personalized data on the internet.

Information, Disclosure, and Preferences. As discussed in the introduction, we impose the key restriction that not all information a firm possesses can be communicated to a given consumer. For simplicity, we suppose that there are two attributes for a given firm's product or service, $i = 1, 2$, and that the firm can only communicate one of these two attributes. Let M represent the set of firms as well as the number of firms. A consumer who knew both attributes of a given firm m would learn two values u_i^m, resulting in the true valuation

$$u^m = \sum_{i=1,2} u_i^m, \tag{1}$$

in the spirit of Lancaster (1966).

Attributes are independently and symmetrically distributed across firms. Also, to preserve symmetry, costs are symmetric and normalized to zero. Ex-ante, for a given consumer each value u_i^m is independently distributed according to $F(u_i^m)$, which is atomless and has everywhere a strictly positive density $f(u_i^m)$. This captures the notion that the characteristics of a given attribute may represent a good match for some but not all consumers. For now we suppose that the support of u_i^m is bounded and given by $[\underline{u}, \overline{u}]$; later we extend our results to allow for unbounded support. Denote $\underline{U} = 2\underline{u}$ and $\overline{U} = E[u] + \overline{u}$.[15]

We consider two communication strategies for firms, depending on whether the single disclosed attribute is chosen selectively or non-selectively. Disclosure is non-selective (or advertising is non-tailored) if consumers know that a firm always discloses the same attribute $d^m \in \{1, 2\}$ to all consumers. Non-selective disclosure results when a firm cannot learn about consumer preferences, so that each consumer looks identical, with match qualities drawn from $F(u_i^m)$. Given symmetry over the two attributes, communicating the same attribute $d^m \in \{1, 2\}$ to all consumers will then be optimal for a firm. For a firm to use selective disclosure vis-à-vis a particular consumer, it must have learned the consumer's preferences,

[15] An alternative interpretation is that disclosure by the firm allows a consumer to learn the distances between the product's true characteristics and her own preferred characteristics. As we show in Appendix C, our key results still hold once we take the distributions of these distances as primitives, while assuming that consumer surplus is reduced by these distances. (This holds despite the fact that the distribution of utility does not necessarily inherit the properties of the distribution of distances.) Using a Salop circle for each attribute, we also derive expressions for a case in which we explicitly disentangle the actual location of product attributes from the preferred location according to the preferences of a particular consumer.

as represented by the values u_1^m and u_2^m. The firm then strategically chooses which attribute d^m to disclose.[16]

Game and Plan of Analysis. We consider the following game. At $t = 1$, firms may or may not learn about the preferences of an individual consumer, depending on whether firms are able to acquire information. This will be the key policy tool. At $t = 2$, firms disclose a particular attribute, d^m, thereby revealing to the consumer the value u_i^m with $i = d^m$. At this stage, they may also engage in personalized pricing when this is feasible, as analyzed in Section 5. At $t = 3$, the consumer decides from which firm to purchase, or whether to purchase at all in the case of monopoly. In the absence of competing firms, with monopoly ($M = 1$) we posit that a consumer's next-best alternative represents an outside option of known value R, for which we stipulate that $\underline{u}+E[u] < R < \overline{u}+E[u]$. This assumption ensures that a consumer will both purchase and not purchase with strictly positive probability. In the case of monopoly, we conveniently drop the firm subscript $m = 1$.

Section 3 lays the foundation for our analysis by comparing consumer updating under selective and non-selective disclosure. Sections 4 and 5 build on these results to analyze the outcomes with and without regulation, depending on whether or not personalized pricing is feasible.

3 Selective versus Non-Selective Disclosure

In this section we focus on how a consumer forms beliefs about the value of the product attributes offered by a particular firm. For now, it is convenient to omit the firm's superscript, m. With non-selective disclosure, a consumer learns the disclosed attribute but retains the initial prior belief about the non-disclosed attribute. After observing the realization u_d of non-selectively disclosed attribute d, the consumer's expected valuation for the product is $U = u_d + E[u]$. From an ex-ante perspective, a consumer's expected valuation under non-selective disclosure then follows the distribution function

$$N(U) = F(U - E[u]), \tag{2}$$

with support between $\underline{u} + E[u]$ and $\overline{u} + E[u]$.

[16]Note that this modeling of selective disclosure entails the restriction that the firm must communicate one of the two attributes, $d^m \in \{1, 2\}$. Hence, here we do not consider the strategy of a firm not to communicate any of the two attributes. When consumers are wary of firms' choices, however, we can show that such a strategy would not arise in equilibrium due to a standard unraveling argument. When consumers are unwary as they naively fail to anticipate firms' disclosure strategies, such a "non-disclosure" strategy, if it is feasible at all, may be sometimes chosen in equilibrium, even though we can still show that our main results extend.

We turn next to the case of selective disclosure. For this suppose that it is known to a consumer that the respective firm always chooses the attribute d with the highest realization u_d. In our subsequent analysis of the full game of disclosure this will indeed be optimal, given that consumers place the same weight on both attributes and given that the "fit" for each attribute is distributed according to the same distribution function $F(u_i)$.[17] When $u_1 = u_2$, which is a zero-probability event, it is likewise known that the firm randomizes. The consumer should thus rationally update that $u_n \leq u_d$ holds for the attribute $n \neq d$ that has not been disclosed, obtaining an overall expected valuation equal to

$$U = u_d + E[u \mid u \leq u_d]. \tag{3}$$

This equation can be solved implicitly for a unique and monotone function $u_S(U)$ that retrieves the value u_d that the consumer must have learned for a given expected value U. Under selective disclosure, the ex-ante distribution of the expected valuation is then obtained using the distribution of the maximum,

$$S(U) = F^2(u_S(U)), \tag{4}$$

with support between $\underline{U}$ and $\overline{U}$.

The conditional expected valuation U in (3) and its distribution (4) were derived under the assumption that the consumer is wary of the fact that the firm selectively communicates one of the two attributes. One reason why this may not be the case is that the consumer is generally not aware of the firm's capability to collect and use personally identifiable data in this way. Equivalently, the consumer may underestimate the skills of a salesperson as well as the underlying conflict of interest. Note that this case will also be relevant off-equilibrium when consumers are wary and when it is not observable that firms collect data and tailor their disclosure accordingly. Then, when faced with u_d, the consumer wrongly believes that her valuation is $\widehat{U} = u_d + E[u]$, thus not discounting the valuation for the adverse selection, as properly done in expression (3). As the firm, however, selectively communicates u_d, in this case the revealed attribute is no longer distributed according to $F(u_d)$, but according to $F^2(u_d)$. The perceived expected valuation for a consumer who remains naively unaware of selective disclosure is then distributed according to

$$S_n(\widehat{U}) = F^2(\widehat{U} - E[u]),$$

with support between $\underline{u} + E[u]$ and $\overline{u} + E[u]$.

[17] In Appendix B, we consider the case where consumers place different weights on different attributes; when F is uniform we obtain a clean characterization and can show that our main comparison of selective and non-selective disclosure still applies.

Impact of Selective Disclosure with Unwary Consumers. Here, the perceived expected valuation under selective disclosure and the true expected valuation under non-selective disclosure have the same support, ranging from $\underline{u} + E[u]$ to $\overline{u} + E[u]$. However, the distribution of the perceived expected valuation $\widehat{U}$ under selective disclosure dominates the distribution of the true expected valuation U under non-selective disclosure in the sense of strict First-Order Stochastic Dominance. For future reference, we state this as follows.

Observation 1 *When a consumer is unwary about selective disclosure, the distribution $S_n(\widehat{U})$ of her perceived valuation $\widehat{U}$ dominates the distribution of her true expected valuation U under non-selective disclosure, $N(U)$, in the sense of strict First-Order Stochastic Dominance.*

As an immediate consequence, note that the ex-ante expected value under selective disclosure to an unwary consumer $E[\widehat{U}]$ strictly exceeds the ex-ante expected value under non-selective disclosure $E[U] = 2E[u]$.

Impact of Selective Disclosure with Wary Consumers. Compared to the distribution with non-selective disclosure, $N(U)$, the distribution of a wary consumer's expected valuation with selective disclosure, $S(U)$, assigns more mass to lower values of U. In fact, the lower bound of the support is no longer $\underline{u} + E[u]$, but $\underline{U} = 2\underline{u}$, which is strictly smaller. Intuitively, when observing $u_d = \underline{u}$ under selective disclosure, a wary consumer correctly updates that the non-disclosed attribute has the worst fit ($u_n = \underline{u}$ with $n \neq d$). We next compare the two distributions at the upper end of their support. Note that this is $\overline{U} = \overline{u} + E[u]$ in either case because, even with selective disclosure, a wary consumer cannot learn about the non-disclosed attribute when the disclosed attribute takes on the highest match value $\overline{u}$. For a given value of the consumer's expected valuation U, let us denote the corresponding disclosed values by $u_N = U - E[u]$ and $u_S = u_S(U)$ (which implicitly solves equation (3)). That is, u_N and u_S are the backed-out values of the disclosed attribute u_d under non-selective and selective disclosure. Recall that at the highest possible value $U = \overline{U}$ we have $u_S = u_N = \overline{u}$. The densities of the two distributions are then

$$\begin{aligned} N'(U) &= \frac{dN(U)}{dU} = f(u_N), \\ S'(U) &= \frac{dS(U)}{dU} = f(u_S)\frac{2F(u_S)}{1 + \frac{dE[u|u \leq u_S]}{du_S}}. \end{aligned}$$

Suppose now that f is logconcave. As is well known, this implies that

$$\frac{dE[u \mid u \leq u_S]}{du_S} < 1. \tag{5}$$

Together with $u_S = u_N$ at the highest realization of the expected valuation $U = \overline{U}$, next to $F(u_S = \overline{u}) = F(u_N = \overline{u}) = 1$, this yields $N'(\overline{U}) < S'(\overline{U})$. Thus, when f is logconcave, with selective disclosure the distribution of a wary consumer's posterior valuation has more mass in the upper tail, compared to the distribution when communication is non-selective: $N(U) > S(U)$ for all sufficiently large values of U. In essence, there are two forces at play here: first, with selective disclosure it is more likely that higher values of u are disclosed; second, this is, however, discounted as the consumer updates about the non-disclosed attribute. Condition (5) ensures that in the upper tail the first effect dominates as, when we decrease u_S slightly compared to $\overline{u}$, the reduction in the conditional expected value of the non-disclosed attribute remains sufficiently bounded. Taken together with our previous observations, when the distribution f is logconcave, selective disclosure to a wary consumer thus results in a distribution of the expected valuation U that has more mass in the lower and upper tail (while clearly not affecting the expected value $E[U] = 2E[u]$). As we show in the proof of Observation 2, when f is logconcave but now with unbounded upper support $\overline{u} = \infty$, there is still more mass in the upper tail with selective disclosure: $N(U) > S(U)$ for all sufficiently large U.

Examples: To go beyond this tail result, let us first take a tractable example with a uniform distribution on $[\underline{u}, \overline{u}]$. As is easily confirmed, we obtain the distributions

$$N(U) = \frac{2U - 3\underline{u} - \overline{u}}{2(\overline{u} - \underline{u})} \text{ and } S(U) = \left(\frac{2(U - 2\underline{u})}{3(\overline{u} - \underline{u})}\right)^2. \tag{6}$$

In this case, there is a single point of intersection in the interior of both supports, namely at $\widetilde{U} = (3\overline{u} + 5\underline{u})/4$. Next, when u_i follows an exponential distribution (whose support is unbounded above), we can show that the single-crossing property also holds generally: $S(U)$ and $N(U)$ intersect exactly once (cf. the proof of Observation 2). Figure 1 illustrates the comparison between $N(U)$, $S(U)$, and $S_n(U)$ for an example where the value distribution is uniform on $[0, 1]$.

We have thus established the single-crossing property for the tractable uniform case and, in the family of distributions with unbounded support, for the case with an exponential distribution.

Observation 2 *When a consumer is wary about selective disclosure, then the distribution $S(U)$ of her expected valuation U and the distribution under non-selective disclosure, $N(U)$, compare as follows. When the attribute value u_i is either uniformly or exponentially distributed, a single-crossing property holds; there is thus a value $\underline{u} + E[u] < \widetilde{U} < \overline{u} + E[u]$ such that, as long as U is in the interior of at least one support, $S(U) > N(U)$ for $U < \widetilde{U}$ and*

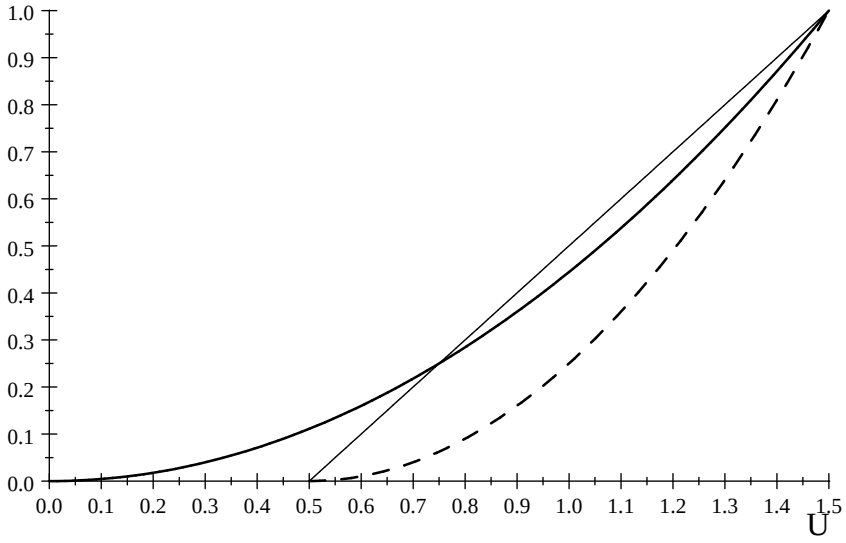

Figure 1: **Comparison of Distributions.** For an example with uniformly distributed attributes, $F(u) = u$, this graph illustrates the general comparison between the ex-ante distributions of the overall utility, U, under non-selective disclosure (regular segment), $N(U)$, selective disclosure to a wary consumer (bold curve), $S(U)$, and selective disclosure to an unwary consumer (dashed curve), $S_n(U)$.

$S(U) < N(U)$ for $U > \widetilde{U}$. Generally, when $f(u_i)$ is logconcave, $S(U)$ always has more mass than $N(U)$ in the lower and upper tail.

Proof. See Appendix A.

We have made many numerical calculations with all typically used logconcave distributions and have always found the single-crossing property between $N(U)$ and $S(U)$ to hold.[18] In what follows, we restrict the analysis to distribution functions $F(u_i)$ that indeed imply such a single-crossing property.[19]

Assumption: *$F(u_i)$ is such that $N(U)$* and *$S(U)$ satisfy the single-crossing property, so that there exists $\widetilde{U}$ with $S(U) > N(U)$ for $U < \widetilde{U}$ and $S(U) < N(U)$ for $U > \widetilde{U}$.*

[18] Precisely, we have experimented with all distributions with logconcave density functions listed in Table 1 of Bagnoli and Bergstrom (2005).

[19] Note also that in contrast to Johnson and Myatt (2006), in our setting we compare two distributions, rather than considering a larger family of distributions, which could arise from a continuous change in one parameter. They focus on a more continuous rotation around the point of intersection (for all distributions in the considered family). To introduce a more continuous version of selective communication in our model, for example, by letting the firm observe only a noisy signal of u_i, we would need to choose a particular functional specification.

We should note that for some results, as will become evident, also a weaker condition would be sufficient, namely that $S(U)$ results from $N(U)$ through a (mean-preserving) spread: $\int_{\underline{U}}^{y}[S(U)-N(U)]dU > 0$ for all $y < \overline{U}$. More precisely, in what follows all comparisons that relate to consumer surplus would go through with this weaker condition. Instead, we need the single-crossing property for some results on firms' optimal choice between selective and non-selective disclosure.

4 Selective Disclosure without Personalized Pricing

This section considers cases where firms cannot use information about individual consumer preferences to price discriminate. Given that each firm offers all consumers the same product, even when it selectively gives them different information, personalized pricing may be impossible or at least difficult with physical goods that can be easily resold. Price discrimination would then create scope for arbitrage, either through a grey (or parallel) market between consumers or through the activity of intermediaries. Also, price discrimination may be limited when consumers are concerned about fairness.[20] In the case of face-to-face interaction with a salesperson, a firm may also be reluctant to grant its agent control over the product's price. Furthermore, when the considered channel may only represent one among several (online or offline) distribution channels, the firm's pricing flexibility for this channel may be seriously compromised, so that we may indeed abstract away from pricing differences depending on the firm's disclosure policy. (See also the application to political campaigning at the end of this section.)

Given that firms are symmetric, when consumers compare the different offers of $M > 1$ firms, abstracting from the common price level, consumers will choose the product (or firm) with the highest expected utility: $U^{(1)} = \max_{m \in M} U^m$. Likewise, denote $\widehat{U}^{(1)} = \max_{m \in M} \widehat{U}^m$ in the case of unwary consumers, whose perceived value may differ from the true expected value. In case of selective disclosure, firms will disclose the highest-value attribute.[21]

In what follows, it proves convenient to first ask about consumers' preferences regarding

[20] Price (or rate) parity has become a major objective for firms, e.g., hotels, given the increasing transparency via online channels.

[21] Precisely, with competition this is uniquely optimal. The case of monopoly ($M = 1$) is slightly different. The consumer then compares the expected utility U (or $\widehat{U}$), dropping the firm superscript, with the outside option's known value, R. (To express results uniformly for all M, the outside option has already been adjusted for by the respective price differential.) When both attributes generate a low or high value, the monopolist is indifferent about which to disclose, given that either attribute will trigger either no purchase or a sure purchase. Even in this case, however, robustness considerations lead us to stipulate that the monopolist will always choose to disclose the highest attribute. In fact, this indifference would no longer prevail if, with arbitrary small probability $\varepsilon > 0$, the consumer were to privately observe an additional component (with sufficiently large support) to her utility from either the product or the outside option.

selective or non-selective disclosure (and the respective data collection that the former requires). This analysis will provide the background to subsequently compare the unregulated market equilibrium with the outcome under regulation.

4.1 Consumer Preferences

Wary Consumers. Denote by M_S the set, as well as the number, of firms that choose selective disclosure, again with a slight abuse of notation. Suppose that this set is known to wary consumers. We establish first that wary consumers are strictly better off when more firms choose selective disclosure, i.e., when M_S becomes larger. Intuitively, this result follows from the fact that more information—corresponding to the fact that, for each product offered by a firm in M_S, the undisclosed attribute has a value lower than the level of the disclosed attribute—becomes available with selective disclosure. Wary consumers are able to use this information to make a better purchase decision.

It is convenient to denote by $G^m(U)$ the (true) distribution of the consumer's expected utility for product m, so that $G^m(U) = N(U)$ when firm m chooses non-selective disclosure and $G^m(U) = S(U)$ when it chooses selective disclosure. Denote by $U^{(1:M\setminus m)}$ a consumer's maximum expected utility over the products of all firms other than m, with corresponding distribution $G^{(1:M\setminus m)}(\cdot)$. The ex-ante expected utility for a wary consumer can then be written as

$$E[U^{(1)}] = \int_{\underline{U}}^{\overline{U}} \left[\int_{\underline{U}}^{\overline{U}} \max\left\{U^{(1:M\setminus m)}, U^m\right\} dG^m(U^m) \right] dG^{(1:M\setminus m)}(U^{(1:M\setminus m)}). \tag{7}$$

Here, we first take the expectation with respect to the distribution of the consumer's highest valuation for all other products and then with respect to the maximum of this and the valuation for product m. We next compare the case where firm m chooses selective disclosure, $G^m(U) = S(U)$, with the case where it chooses non-selective disclosure, $G^m(U) = N(U)$. Making use of (7), the difference of the consumer's respective ex-ante expected utility, after integration by parts, is

$$\int_{\underline{U}}^{\overline{U}} \left[\int_{U^{(1:M\setminus m)}}^{\overline{U}} \left[N(U^m) - S(U^m)\right] dU^m \right] dG^{(1:M\setminus m)}(U^{(1:M\setminus m)}). \tag{8}$$

From the single-crossing property of $N(U)$ and $S(U)$, the integral in brackets in (8) is strictly positive for all interior $U^{(1:M\setminus m)}$, so that the total integral is strictly positive.[22] A wary consumer thus indeed strictly benefits when firm m applies selective disclosure, regardless

[22] Clearly, for this result, the weaker requirement of a mean-preserving spread would have also been sufficient.

of whether competitors choose selective or non-selective disclosure. Note further that the preceding argument applies also to the case of monopoly. There, recall that a consumer compares the expected valuation U for the monopolist's product with the valuation R for an outside option. Then, expression (8) boils down to

$$\int_{R}^{\overline{U}} [N(U) - S(U)]\, dU. \tag{9}$$

Again, expression (9) is strictly positive from the single-crossing property of $N(U)$ and $S(U)$.

Proposition 1 (Welfare of Wary Consumers) *When consumers are wary, a known switch of any firm from non-selective to selective disclosure strictly benefits consumers. Consumer welfare is thus maximized when all firms choose selective disclosure.*

Unwary Consumers. For the case with unwary consumers it makes a difference whether there is competition rather than monopoly, as well as whether all or only a subset of firms practice selective disclosure in the case of competition. Consider first the case where there are two or more firms that compete and they all choose selective disclosure, $M_S = M > 1$. Recall that an unwary consumer overestimates her expected utility as she does not adjust downwards her expectation for the "fit" of the non-disclosed attribute: $\widehat{U}^m > U^m$. When all firms choose selective disclosure, however, by symmetry this bias equally affects the consumer's expectation for all firms' products. In fact, an unwary consumer's decision rule is then the same as that of a wary consumer, namely to simply purchase from the firm where the respective disclosed value u_i^m with $i = d^m$ is maximal. Thus, it follows that, when all firms choose selective disclosure, unwary consumers realize the same expected utility that wary consumers realize. Given that unwary consumers clearly realize the same expected utility as wary consumers also when firms choose non-selective disclosure, we conclude from Proposition 1 that unwary consumers, like wary consumers, are strictly better off with selective disclosure.[23] Thus, competition protects unwary customers from the adverse selection generated by selective disclosure, given that all firms practice it symmetrically. This simple and powerful insight seems novel to the literature.

As we show below, this comparison between the two extreme cases, where all or none of the firms practice selective disclosure, will be particularly relevant when consumers are unwary, because then all firms strictly prefer selective disclosure, regardless of other firms'

[23] When only some firms choose selective disclosure, it is generally ambiguous how one firm's choice affects the ex-ante true expected utility of naive consumers. When this firm switches to selective disclosure, the switch increases the ex-ante likelihood that the firm's product is purchased, which creates a bias vis-à-vis products of firms with non-selective disclosure but corrects for a bias vis-à-vis products of firms with selective disclosure. This is analoguous to the case of a monopoly that we analyze below.

choices, while the outcome with non-selective disclosure may, instead, be prescribed by regulation.

The case of monopoly is special, however. When there is only one firm that can choose selective disclosure, the consumer compares the resulting expected value with the known value from an outside option. Thus, selective disclosure unambiguously biases an unwary consumer's relative perception of the monopolist's product vis-à-vis the outside option. Still, an unwary consumer may be better off. To see this, take the two types of errors that consumers can make. The mistake of erroneously making a purchase even though $u_1+u_2 < R$ evidently becomes larger when an unwary consumer faces selective disclosure. This is an immediate implication of the observation that with selective disclosure, an unwary consumer always ends up purchasing whenever he would do so with non-selective disclosure. On the other hand, it is also less likely that a consumer does not purchase, even when $u_1 + u_2 > R$. How these two errors trade off should generally depend on the distribution $F(u_i)$. For the special case with a uniform distribution, we find that the unwary consumer's true expected utility is exactly the same under the two regimes. In the example with an exponential distribution, unwary consumers are in fact strictly better off with selective disclosure, even though in the case of monopoly their decision is biased.

Proposition 2 (Welfare of Unwary Consumers) *With competition, unwary consumers are strictly better off when all firms choose selective disclosure, compared to the case where all firms choose non-selective disclosure. In the case of monopoly, the comparison is generally ambiguous; unwary consumers' true expected utility is the same under selective and non-selective disclosure if u_i is uniformly distributed, while it is strictly higher under selective disclosure if u_i is exponentially distributed, in spite of the bias generated in the value perceived by consumers.*

Proof. See Appendix A.

4.2 Firm Preferences and Policy

To analyze the potential impact of regulation, we first need to study firms' preferences and the resulting equilibrium without regulation. Our first result is the following. If firms' choice whether or not to collect customer-specific information in $t = 1$ is unobservable, in equilibrium each firm will choose selective disclosure. This holds irrespective of consumers' wariness, precisely as a switch from non-selective to selective disclosure is then not observable. Recall that when a consumer does not anticipate that data is collected and used for selective disclosure, this results in a first-order stochastic upward shift of the distribution of the *perceived* utility (cf. Observation 1). The same logic clearly applies also when consumers

are unwary, in this case regardless of whether the firm's choice to collect information and practice selective disclosure is observable or not.

Proposition 3 (Firms' Choice with Unwary Consumers) *When consumers are either unwary or they are unable to observe the firms' choice to collect personal data and practice selective disclosure, in equilibrium all firms choose selective disclosure.*

The case where consumers are wary and firms' choice is observable is more subtle. Consider the case of a monopolist who could commit not to collect (and therefore not to use) consumer-specific data. Suppose first that the monopolist, however, does not make use of this commitment. Then, from an ex-ante perspective, a purchase will take place with probability $1 - S(R)$. Otherwise, i.e., when the monopolist commits not to practice selective disclosure, the respective probability is $1 - N(R)$. Given single crossing, the respective probability is strictly higher with selective disclosure if R lies to the right of $\widetilde{U}$, while otherwise it is strictly lower. As selective disclosure puts more mass into the tails of the distribution of wary consumers' expected utility, the monopolist's profits increase only if *a priori* a purchase is not too likely.[24] Otherwise, when *a priori* a purchase from the monopolist is sufficiently likely, the monopolist profits from committing to keep disclosure non-selective.

With competition, when the number of firms becomes sufficiently large, we obtain the clear-cut result that all firms will choose selective disclosure. Intuitively, in this case, regardless of other firms' choices, the distribution of the first-order statistic over the expected utility from all other firms, $U^{(1:M\setminus m)}$, has increasingly more mass in the right-hand tail. What then matters for the choice of a single firm's disclosure strategy is the likelihood that a consumer's expected utility is very high. From the single-crossing property, selective disclosure assigns more mass to the right-hand tail of a consumer's expected valuation.[25]

Proposition 4 (Firms' Choice with Wary Consumers) *When consumers are wary and they are able to observe whether firms choose to collect personal data so as to practice selective disclosure, the equilibrium choice of disclosure by firms depends on the intensity of competition:*
(i) A monopolist ($M = 1$) strictly prefers selective disclosure if a consumer is a priori unlikely to purchase, as R exceeds $\widetilde{U}$, but prefers non-selective disclosure if, instead, $R < \widetilde{U}$.

[24] This is thus akin to Johnson and Myatt's (2006) "niche market".

[25] Generally, when M is still relatively small, we cannot determine unambiguously firm preferences. To see this, the analogy to the monopoly case is instructive. From the perspective of firm m we can treat $U^{(1:M\setminus m)}$ in analogy to the reservation value R in the case of monopoly. Again, for given $U^{(1:M\setminus m)}$ firm m would want selective disclosure if and only if this is relatively high. Whether the firm thus prefers selective disclosure from an ex-ante perspective hinges, amongst other things, on the distribution of $U^{(1:M\setminus m)}$. When we turn to the application to political campaigning, we will establish as a side result that, with uniform distribution and $M = 2$, a firm strictly prefers selective disclosure regardless of its rival's choice.

(ii) When the number M of competing firms is sufficiently large, all firms choose selective disclosure ($M_S = M$).

Proof. It remains to deal with assertion (ii) for the case with $M > 1$. Recall that we denote the distribution of $U^{(1:M\setminus m)}$ by $G^{(1:M\setminus m)}(\cdot)$, while U^m is distributed according to $G^m(\cdot)$, with either $G^m(U) = N(U)$ or $G^m(U) = S(U)$. From an ex-ante perspective, the likelihood with which a wary consumer buys from firm m is thus

$$\int_{\underline{U}}^{\overline{U}} [1 - G^m(U)]\, dG^{(1:M\setminus m)}(U),$$

so that selective disclosure dominates non-selective disclosure when

$$\int_{\underline{U}}^{\overline{U}} [N(U) - S(U)]\, dG^{(1:M\setminus m)}(U) > 0. \tag{10}$$

When m_S of all other firms choose selective disclosure, we have

$$G^{(1:M\setminus m)}(U) = S^{m_S}(U)N^{M-m_S-1}(U).$$

As $\int_{\underline{U}}^{\overline{U}} [N(U) - S(U)]\, dU = 0$, we thus have for any given cutoff $\widetilde{U} < \overline{U}$, for which $N(U) - S(U) > 0$ for $U > \widetilde{U}$, that (10) indeed holds whenever M is sufficiently large, irrespective of the corresponding choice of $0 \leq m_S \leq M - 1$. **Q.E.D.**

Propositions 3 and 4 characterize the outcome without regulation, so that firms are free to collect and use personalized data. We next turn to the implications of different forms of regulation aimed at restricting firms' options. Here, we distinguish again between the impact when consumers are wary and when they are unwary.

Regulation with Wary Consumers. We proceed by presuming that, in the absence of regulation, consumers do not observe whether firms collect and use personalized data to practice selective disclosure. Consequently, by Proposition 3, all firms would indeed choose selective disclosure, which is also the best outcome for wary consumers by Proposition 1. In other words, wary consumers strictly benefit when firms learn about consumer preferences and tailor their communication strategy accordingly, thereby revealing the attribute with the best fit. In this sense, we can conclude that selective disclosure by all firms, which is then the equilibrium outcome without regulation, is more informative for wary consumers. Wary consumers are hurt by regulation that prohibits firms from collecting and using personal information for selective disclosure.

Interestingly, even a less restrictive regulation requiring firms to seek consumer consent can backfire and lead to a reduction in consumer surplus. This is the case when firms would

like to commit not to engage in selective disclosure, but cannot do so as their strategy to collect and use personalized data is not transparent to consumers. Regulation that prescribes consumer consent then provides such commitment, which is in the interest of firms but not in the interest of consumers. We know from Proposition 4 that a monopolist would strictly benefit from such seemingly light-touch regulation when it is a priori already sufficiently likely that consumers will purchase under non-selective disclosure ($R < \widetilde{U}$). When, instead, many firms (large M) vigorously compete, we know from Proposition 4 that they will all profit from selective disclosure and will thus seek consumer consent, so that regulation that merely requires consumer consent would have no consequences.

Proposition 5 (Regulation with Wary Consumers) *Regulation that restricts selective disclosure, for example, by preventing the collection of personal data, always harms wary consumers. Regulation that requires firms to obtain consumer consent is harmful when it enables firms to commit vis-à-vis wary consumers not to practice selective disclosure, and it is inconsequential at best.*

Regulation with Unwary Consumers. We know from Proposition 3 that, when facing unwary consumers, a firm would always want to gather customer-specific information. Suppose that consumers remain unaware about the practice of selective disclosure even when a firm must obtain their consent to acquire personal information. This may be the case, for instance, when an internet platform only asks once for general permission to collect data, so that the specific consequences of this consent may not be salient for the consumer at the time he sees a specific marketing campaign. Thus, in this case the requirement to obtain consent from consumers does not constrain firms, once we ignore the costs of obtaining such consent.

What are then the implications for unwary consumers of an outright ban on collecting and using personalized data? From the preceding observations, we have to compare only the cases where either all firms (no regulation) or none of them (regulation) practice selective disclosure. When $M > 1$ firms compete, by Proposition 2 we know that unwary consumers are strictly worse off with a ban. The case of monopoly is different, because consumers are no longer protected from the bias generated by selective disclosure when they compare the monopolist's offer to the known value of an outside option rather than to the offer of a competing firm. However, recall from Proposition 2 that despite this bias, unwary consumers may still (unknowingly) benefit from selective disclosure because of a reduction in the mistake of erroneously not purchasing even though $u_1 + u_2 > R$.

Proposition 6 (Regulation with Unwary Consumers) *When consumers are unwary about selective disclosure and remain so after giving their general consent to the collection*

of their personalized data, regulation that requires firms to obtain such consent is inconsequential. An outright ban prohibiting information collection or use hurts unwary consumers when there are competing sellers ($M > 1$). With a monopolist seller ($M = 1$), the outright ban has an ambiguous effect on consumers in general; true consumer surplus remains unaffected by regulation if u_i is uniformly distributed and is strictly reduced if u_i is exponentially distributed.

4.3 Hypertargeting Political Campaigns

> "Political campaigns, which have borrowed tricks from Madison Avenue for decades, are now fully engaged on the latest technological frontier in advertising: aiming specific ads at potential supporters based on where they live, the Web sites they visit and their voting records. . . .
>
> The process for targeting a user with political messages takes three steps. The first two are common to any online marketing: a "cookie," or digital marker, is dropped on a user's computer after the user visits a Web site or makes a purchase, and that profile is matched with offline data like what charities a person supports, what type of credit card a person has and what type of car he or she drives. The political consultants then take a third step and match that data with voting records, including party registration and how often the person has voted in past election cycles, but not whom that person voted for.
>
> Throughout the process, the targeted consumers are tagged with an alphanumeric code, removing their names and making the data anonymous. So while the campaigns are not aiming at consumers by name—only by the code—the effect is the same. Campaigns are able to aim at specific possible voters across the Web." *Tanzina Vega, Online Data Helping Campaigns Customize Ads, New York Times, February 20, 2012.*

Extending the Model. We now extend the model to political campaigning. Suppose for specificity that there is competition for voters by $M = 2$ candidates. Candidates' platforms comprise two issues $i = 1, 2$, along which a candidate's stance can more or less coincide with the preference and political orientation of a particular voter. Equivalently, the model also captures voting for or against a motion brought forward in a debate or referendum.

For convenience only, suppose there is an odd number $V \geq 3$ of voters; the case with $V = 1$ coincides with our previous analysis of consumer choice. Denote by $u_i^m(v)$ voter v's independent draw on issue i and candidate m.[26] Each voter casts the ballot for the candidate whose future decisions promise to deliver the highest utility value, (1).[27] The candidate m

[26] See again also Appendix C, where we separately model a sender's location and a receiver's preferred location, each along a Salop circle, and express the receiver's valuation as a function of the perceived difference between the actual and the preferred location.

[27] See, however, Appendix B, where we allow the different issues/attributes to carry different weights.

who obtains the highest number of voters wins the election, resulting in the overall value $u^m(v)$ for voter v. Denote by $U^m(v)$ a voter's expected utility when candidate (or motion) m wins and by $\widehat{U}^m(v)$ the respective perceived utility, which can differ when voters are unwary about selective communication.

Wary Voters. Using ex-ante symmetry across voters, the likelihood of a vote for candidate $m = 1$, corresponding to $U^1(v) \geq U^2(v)$, is given by

$$q = \int_{\underline{U}}^{\overline{U}} \left[1 - G^{m=1}(U)\right] dG^{m=2}(U), \tag{11}$$

where the tie-breaking assumption is clearly without loss of generality given that ties happen with zero probability. When V^m is the number of votes cast for candidate m out of a total of V votes, the ex-ante probability with which candidate $m = 1$ is elected is then $Q = 1 - \Pr[V^1 \leq (V-1)/2]$. Conveniently, given that V is odd, this becomes

$$Q = 1 - \sum_{k=0}^{\frac{V-1}{2}} \binom{V}{k} q^k (1-q)^{V-k},$$

which is clearly strictly increasing in q. The likelihood $1 - Q$ with which candidate $m = 2$ is elected is strictly decreasing in q. Further, let y be the (symmetric) probability with which a given voter will become pivotal. Using again that V is odd, we have

$$y = \binom{V-1}{\frac{V-1}{2}} q^{\frac{V-1}{2}} (1-q)^{\frac{V-1}{2}}. \tag{12}$$

For each voter, his ex-ante utility now comprises two terms. If the voter ends up not being pivotal, the vote does not influence the decision; given that preferences are independently drawn, the expected utility equals the unconditional expectation $2E[u]$. Instead, when this voter is pivotal, the conditional expected utility equals $E[U^{(1)}]$. Multiplied with the respective probabilities, y for becoming pivotal and $1 - y$ otherwise, a voter's ex-ante expected utility is

$$2E[u] + y\left\{E[U^{(1)}] - 2E[u]\right\}. \tag{13}$$

From Proposition 1 we already know that the term in braces is strictly higher when one candidate starts practicing selective disclosure and even higher when both candidates disclose selectively. It is in this sense that our previous analysis can be applied. However, in this application to elections, the decision is no longer determined by the preferences of each individual consumer alone, as in our baseline application to product marketing, but instead by the aggregate preferences of all voters. Therefore, we have to take into account a second effect. The communication strategies of candidates now also affect a voter's likelihood

y of becoming pivotal and, thereby, affect voter utility according to expression (13). The likelihood of becoming pivotal is clearly highest when $q = 1/2$, which applies when both candidates choose the same communication strategy. Instead, when one candidate chooses a different communication strategy, $q \neq 1/2$ can result, so that y decreases. In fact, this is always the case when one candidate alone, say $m = 1$, chooses to hypertarget and strictly prefers to do so, as then $q > 1/2$. In this case, a trade-off results given that the improved informativeness increases a voter's utility conditional on being pivotal, while the probability that any given voter becomes pivotal decreases.

Before we explore this trade-off further, note that such an asymmetric situation may arise when only one candidate's campaign is sufficiently sophisticated. Furthermore, with a uniform distribution we can show that a candidate indeed strictly prefers hypertargeting as $q > 1/2$ when only candidate 1 hypertargets or, likewise, $q < 1/2$ when this is done so only by candidate 2 (cf. Proposition 7).[28] We show next how generally the trade-off mentioned above is resolved unambiguously when the number of voters is sufficiently high. For this, denote by q_S and y_S a voter's probability of voting for 1 and for being pivotal when only candidate 1 hypertargets, and by y_N and $q_N = 1/2$ the corresponding probabilities when both candidates disclose non-selectively. With a slight abuse of notation, $E_S[U^{(1)}]$ and $E_N[U^{(1)}]$ denote the respective conditional utilities of a pivotal voter. Then, a voter's ex-ante expected utility, as given by (13), is higher when candidate 1 hypertargets if

$$\frac{y_S}{y_N} = [4q_S(1-q_S)]^{\frac{V-1}{2}} > \frac{E_N[U^{(1)}] - 2E[u]}{E_S[U^{(1)}] - 2E[u]}. \tag{14}$$

While the right-hand side of the inequality (14) does not depend on V (and is strictly smaller than one), the left-hand side is strictly decreasing in V (and goes to zero) as long as $q_S \neq 1/2$. Hence, either condition (14) does not hold for all V, including $V = 3$, or there exists a cutoff value $\widetilde{V}$ such that it is only for $V < \widetilde{V}$ that a voter is better off when (only) candidate 1 chooses selective communication. While for each voter the benefits conditional on being pivotal remain the same when the number of voters increases, this makes it relatively less likely that a voter becomes pivotal when candidates choose different communication strategies (even though, as noted above, both y_S and y_N approach zero as $V \to \infty$).

Proposition 7 (Welfare of Wary Voters) *Suppose V voters decide by majority rule over $M = 2$ candidates and that voters are fully aware of the occurrence of hypertargeting. Then, voter expected utility is maximized if both candidates hypertarget. A trade-off arises when only*

[28] This result with a uniform distribution hinges also on the fact that both candidates are ex-ante symmetric. In fact, we can show that when one candidate is a priori sufficiently more likely to be selected, given that the support of the respective uniform distribution of match preferences shifts sufficiently upwards, then this candidate no longer prefers selective disclosure.

one candidate hypertargets: asymmetric hypertargeting still benefits each voter conditional on being pivotal, but decreases the likelihood with which individual voters become pivotal in the first place and, thereby, can decide the outcome based on their own preferences. Voters are then either always strictly worse off when only one candidate hypertargets or, as is the case with a uniform distribution over preferences, there exists a cutoff $\widetilde{V}$ on the number of voters such that they are worse off when $V > \widetilde{V}$ and better off when, instead, $V \leq \widetilde{V}$.

Proof. See Appendix A.

Unwary Voters. For a voter who is unaware of the occurrence of hypertargeting, the perceived preference for the respective candidate deviates from their true preference. Using the same notation as in our baseline analysis, voter v will then vote for $m = 1$ when $\widehat{U}^1(v) \geq \widehat{U}^2(v)$.

To begin, suppose that all voters are unwary. Denote the probability that $m = 1$ is elected by $\widehat{q}$, where $\widehat{q} = 1/2$ holds when both candidates choose the same communication strategy. When voters are unwary, $\widehat{q} > 1/2$ holds when only candidate 1 hypertargets and candidate 2 does not. In fact, it is immediate that in this case $\widehat{q} = 2/3$.[29]

The probability $\widehat{y}$ that any (unwary) voter becomes pivotal is again maximized when both candidates choose the same communication strategy, so that $\widehat{q} = 1/2$. When only one candidate hypertargets, the reduction in the likelihood of becoming pivotal has a negative effect on voter welfare. With wary voters, however, we showed for the tractable case of a uniform distribution that for sufficiently low V this is more than compensated by the higher conditional expected utility when a voter becomes pivotal, given the higher informativeness of selective disclosure. This is no longer the case, instead, when voters are unwary. In fact, for the case of a uniform distribution, we argue next that an unwary voter's true expected utility conditional on being pivotal when only one candidate hypertargets is the same as when both candidates communicate non-selectively, so that, independent of the number of voters, only the first, negative effect survives.

With a uniform distribution, the argument why, more generally, an unwary decision-maker, i.e., a pivotal voter or a consumer in our baseline application, realizes the same true expected utility when none or exactly one candidate (or firm) chooses selective disclosure is as follows. Recall from Proposition 2 for the case of monopoly with uniformly distributed

[29] We can transform $\widehat{q} = \int_{\underline{U}}^{\overline{U}} [1 - S_n(U)]\, dN(U)$ into

$$\int_{\underline{u}+E[u]}^{\overline{u}+E[u]} 2f(U - E[u])F^2(U - E[u])dU = \frac{2}{3}F^3(U - E[u])\Big|_{\underline{u}+E[u]}^{\overline{u}+E[u]} = \frac{2}{3}.$$

valuations, the true expected payoff of an unwary consumer is the same regardless of the occurrence of hypertargeting—because then the improved information exactly offsets the increased bias. This result was shown to hold for any choice of the reservation value R, and thus extends to the case where R is a priori uncertain. With competition between candidates or firms, the (ex-ante uncertain) expected value of the respective alternative choice is now akin to the reservation value R and it is also not biased because, by assumption, there is non-selective disclosure by the other candidate.

Proposition 8 (Welfare of Unwary Voters) *Suppose that V voters decide by majority rule over $M = 2$ candidates. If all voters are unwary of the occurrence of hypertargeting, a candidate would always want to hypertarget, and voters are best off when both candidates hypertarget. But voters may be strictly worse off, independent of the number of voters, when only one candidate chooses selective disclosure; this is surely the case when preferences are uniformly distributed.*

Proof. See Appendix A.

When some but not all of the V voters are unwary, by our previous observations the outcome is not affected when both candidates choose either selective or non-selective disclosure. This is no longer the case if instead candidates choose different communication strategies. Then, when V_u of the voters are unwary, the higher is V_u the more likely it is that the hypertargeting candidate will be elected. In the uniform case, where we know that selective communication leads also to a higher likelihood q with which a wary voter elects the respective candidate, an increase in V_u reduces the probability that each voter will become pivotal.

Proposition 9 (Welfare of Mixed Voter Audience) *Suppose V_u voters remain unwary of hypertargeting. When both candidates choose the same communication strategy, the outcome does not depend on V_u. When, instead, only one candidate hypertargets, if this at least weakly increases the likelihood of being elected by each voter (i.e., $q \geq 1/2$, which holds strictly with a uniform distribution), an increase in the number of unwary voters V_u decreases each voter's probability of becoming pivotal and thus also reduces true voter expected welfare.*

Proof. See Appendix A.

Groups of Voters. Our application to voting can apply to various settings, from small committees to larger elections, and the vote could be cast for individual candidates as well as for or against a particular motion. In the application thus far, V has represented the number

of voters. For each voter, the respective fit with a candidate's orientation (or the content of a motion) was chosen independently. Such preferences may also be shared across different voters in an electorate. One way to extend our results is to now suppose that there are V voter groups, each composed of z_v voters with the same preferences or political orientation. A campaign in this case would target groups of voters rather than individual voters.[30] In an increasingly fragmented media landscape this could be achieved by tailoring the campaign message to different channels that are frequented by voters with a particular orientation. Our previous analysis immediately extends to the case in which each of these V groups has identical size $z_v = z$.[31]

5 Personalized Pricing

So far our approach to disclosure to consumers has considered the case of (physical) products for which firms are unable to price discriminate, either because consumers or intermediaries can arbitrage away price differences or because firms fear reputational repercussions. Because of transaction costs, however, arbitrage is not a real concern in many markets for retail products.[32] In addition, first-degree price discrimination may be more easily accomplished and remain less transparent to consumers when selling services. This section turns to situations in which firms are not only able to learn about the preferences of consumers and target their communication accordingly, but are also able to charge personalized prices to customers.

The industrial organization literature on behavior-based price discrimination has focused on personalized pricing where, in particular, the past purchasing history of consumers is used; see, for example, Villas-Boas (1999) and Acquisti and Varian (2005). Taylor (2004), for instance, explicitly considers naive consumers who fail to anticipate that firms base future pricing decisions on past purchases. As we abstract from this dynamic feature, our analysis will be quite different. What will again prove key in our comparison of wary and unwary consumers is that the latter fail to anticipate firms' incentives to collect and use personal information so as to tailor advertising.

[30] Indeed, candidates traditionally make highly targeted speeches at private events, such as the behind-closed-doors fundraiser where shortly before the 2012 election Mitt Romney was unwittingly recorded suggesting that some 47% of Americans are government-dependent "victims" who do not pay taxes or take responsibility for their lives, and about whom "it's not my job to worry".

[31] As suggested by Mitt Romney's gaffe, an important drawback of targeted campaigns is the risk that voters exchange information about the different messages they receive. We leave this extension to future research; see, for example, Galeotti and Mattozzi (2012) for a model in which information sharing among voters reduces the incentives for information disclosure by candidates.

[32] For example, consumers at grocery stores are given discount coupons on the back of the receipt based on their purchases. Catalina Marketing, the global leader of precision marketing, also delivers health-related information to consumers based on the prescriptions they pick up at pharmacies.

With competition, we stipulate that firms learn the utility that the consumer perceives for each product, for example, on the basis of some commonly collected information. When no firm chooses weakly dominated prices, this ensures that, first, the consumer purchases the product with the highest expected utility U^m (or, perceived utility $\widehat{U}^m$, in case this differs for unwary consumers), and that, second, the price that the consumer pays is equal to the second-highest such value. Consequently, a consumer realizes the second-order statistic, denoted by $U^{(2)}$ (or, respectively, $\widehat{U}^{(2)}$). We first establish that with personalized pricing all firms choose selective disclosure in equilibrium, provided that they are free to collect information. This holds irrespective of whether consumers are wary or unwary.

Proposition 10 (Firms' Choice with Personalized Pricing) *With personalized pricing, all firms strictly prefer selective disclosure, so that $M_S = M$ is the unique equilibrium when firms are allowed to collect personalized information.*

Proof. Again, it is immediate that a firm strictly benefits from selective disclosure when consumers are unwary or when they do not observe the choice of disclosure regime. With personalized pricing, we now establish that a firm strictly prefers selective disclosure also when wary consumers observe its occurrence. Recall our notation $U^{(1:M\setminus m)}$ for the highest expected utility over all other $M\setminus m$ firms. Then, the profit of firm m is given by

$$\int_{\underline{U}}^{\overline{U}} \left[\int_{\underline{U}}^{\overline{U}} \max\left\{U^m - U^{(1:M\setminus m)}, 0\right\} dG^{(1:M\setminus m)}(U^{(1:M\setminus m)}) \right] dG^m(U^m).$$

We show that this profit is strictly higher when $G^m(U) = S(U)$ rather than $G^m(U) = N(U)$. To see this, note that the term in braces is strictly convex as a function of U^m (unless $U^m = \underline{U}$ or $U^m = \overline{U}$, where it is linear). In fact, the second-order derivative is just $g^{(1:M\setminus m)}(U^m)$. By the single-crossing property, $S(U)$ originates from $N(U)$ through a mean-preserving spread, which implies that the expected value of a convex function of U is strictly higher under $S(U)$ than under $N(U)$. **Q.E.D.**

Note that we now have the clear-cut result that a firm always strictly prefers selective disclosure, even when this is observed by wary consumers and irrespective of other firms' choices. With personalized pricing, a firm that offers a consumer's preferred choice and can thus make a profit wants to maximize the distance between the consumer's expected utility for its own product and the product of its closest rival, because the firm extracts exactly this difference. Selective disclosure serves this purpose by transferring probability to values in the tails—and in particular in the upper tail—of the distribution of consumers' expected utility.

Wary Consumers. When consumers are wary, the effect of firms' disclosure policy on consumer surplus depends crucially on the number of firms M. In the absence of competition ($M = 1$), it is immediate that the monopolist extracts the entire surplus expected by the consumer, both with and without selective disclosure. In this case, the efficiency gains from selective disclosure (cf. Proposition 1) are fully pocketed by the monopolistic firm, while consumers are indifferent between the two disclosure regimes. When there is competition among $M > 1$ sellers, with personalized pricing wary consumers may be strictly worse off under selective disclosure. From an ex-ante perspective, selective disclosure increases the perceived differentiation and thus pushes up prices to the detriment of consumers.

This downside of selective disclosure and increased dispersion is most evident in the case of duopoly ($M = 2$). While without personalized pricing a consumer realized the maximum of the two expected utilities $U^{(1)} = \max\{U^1, U^2\}$, with personalized pricing the consumer realizes the second-highest value, which is now the minimum $U^{(2)} = \min\{U^1, U^2\}$. The outcome without any restrictions, which from Proposition 10 is selective disclosure by both firms, is now even worse for wary consumers than the outcome where firms cannot collect the necessary information and, therefore, must communicate non-selectively.

However, as the number of firms M becomes sufficiently large, so that it becomes increasingly likely that each firm has a close competitor, the benefits of the improved information obtained through selective communication prevail for consumers. Hence, even with personalized pricing, selective disclosure is beneficial to wary consumers, at least when there is sufficient competition. More precisely, Proposition 11 establishes that regardless of what all other firms do, a wary consumer strictly benefits when a particular firm switches from non-selective to selective disclosure. To show this, we derive the following expression. Pick any firm m with U^m distributed according to $G(\cdot)$ and recall that we denote the maximum of expected utilities over all other firms by $U^{(1:M\setminus m)}$ and, likewise, the second-highest realization by $U^{(2:M\setminus m)}$. The notation for the respective distributions then is $G^{(1:M\setminus m)}(\cdot)$ and $G^{(2:M\setminus m)}(\cdot)$. Next, we define the difference in the distributions of the utility for any given firm under non-selective and selective disclosure by $z(U) := N(U) - S(U)$. Then, we show that the consumer benefits when this firm m switches to selective disclosure if and only if

$$\int_{\underline{U}}^{\overline{U}} z(U) \left[G^{(2:M\setminus m)}(U) - G^{(1:M\setminus m)}(U)\right] dU > 0. \tag{15}$$

Given our single-crossing condition for $N(U)$ and $S(U)$, in Proposition 11 we can sign expression (15) unambiguously to be positive whenever there are sufficiently many firms. Then, it always holds that the positive information effect of selective disclosure dominates the negative differentiation effect under personalized pricing.

Proposition 11 (Welfare of Wary Consumers with Personalized Pricing) *With personalized pricing, wary consumers are indifferent between selective and non-selective disclosure by a monopolist ($M = 1$). With duopoly ($M = 2$), consumers are always strictly worse off when a firm m switches to selective disclosure, regardless of the disclosure strategy of the rival firm. However, irrespective of the other firms' choices, consumers benefit when an individual firm chooses selective disclosure provided that there is sufficient competition (M large).*

Proof. See Appendix A.

Proposition 11 relates our paper to results by Board (2009) and Ganuza and Penalva (2010) on the effect of providing bidders with private information in a private-values second-price auction. With personalized pricing, a comparison of consumer surplus in our model effectively amounts to comparing the expectation of the second-order statistic $E[U^{(2)}]$, as in a second-price auction. In these papers, however, the question that is asked is whether providing more information to all bidders increases the auctioneer's expected payoff, while for Proposition 11 we ask whether the switch to selective disclosure by a single firm benefits the consumer.

Unwary Consumers. Recall now that when $M > 1$ and all firms choose selective disclosure, unwary consumers end up purchasing the same product as wary consumers; namely, they all purchase the product with the highest fit. However, once we allow for personalized pricing, there is an interesting difference between the cases with wary and unwary consumers because prices are different in the two cases. What is more, while with selective disclosure a monopolist can charge unwary consumers a strictly higher price than wary consumers, under competition firms can instead charge strictly higher prices to wary consumers.

To see this, take first the case of competition, $M > 1$. The true expected consumer surplus, for given realizations of u_i^m, equals $U^{(2)}$. As we have set firms' costs to zero, a wary consumer buys from the firm disclosing the highest "fit", which generates the expected utility $U^{(1)}$, and pays the price $U^{(1)} - U^{(2)}$. Instead, while an unwary consumer still buys from the firm disclosing the highest fit, so that the true expected utility is $U^{(1)}$, the price that an unwary consumer pays equals $\widehat{U}^{(1)} - \widehat{U}^{(2)}$. If $u^{(1)}$ is the highest disclosed fit and $u^{(2)}$ the second-highest, an unwary consumer pays the price $\widehat{p} = u^{(1)} - u^{(2)}$, given that the expectations about the non-disclosed attribute of either firm wrongly remain unchanged at $E[u]$. A wary consumer pays, instead, the strictly higher price

$$p = u^{(1)} - u^{(2)} + \left\{E[u \mid u \leq u^{(1)}] - E[u \mid u \leq u^{(2)}]\right\},$$

given that the term in braces, equal to the difference in the updated conditional expectations for the non-disclosed attributes by firms 1 and 2, is strictly positive. In other words, in the eyes of an unwary consumer, the firms offering the first-best and the second-best fit appear to be less differentiated, compared to the perceptions of wary consumers. This allows the winning firm to charge only a strictly lower price. Consequently, we can already conclude that, when $M > 1$ and when wary consumers are better off under selective disclosure, unwary consumers are surely strictly better off under selective disclosure.

In the case of monopoly ($M = 1$), however, the comparison with the outcome with wary consumers under selective disclosure is strikingly different. When competition no longer protects unwary consumers, a monopolist can exploit their inflated perceptions, $\widehat{U} > U$. In fact, as the monopolist extracts a price equal to the perceived valuation $\widehat{U}$, and as this is strictly above the true valuation U (unless $u_d = \overline{u}$, which happens with zero probability), the true consumer surplus net of the price is strictly below what the consumer could obtain from staying out of the market. In this sense, an unwary consumer is truly exploited when a monopolist practices selective disclosure and personalized pricing.

Proposition 12 (Welfare of Unwary Consumers with Personalized Pricing) *With personalized pricing and unwary consumers, selective disclosure by all firms in the market has the following impact on the expected utility of consumers: In the case of monopoly ($M = 1$), unwary consumers are strictly worse off than with non-selective disclosure and, in fact, are exploited to the extent that they would be strictly better off when staying out of the market. When there is competition ($M > 1$), unwary consumers are strictly better off than wary consumers under selective disclosure and personalized pricing, implying that they strictly benefit from selective disclosure at least when this holds weakly for wary consumers.*

Regulation. From Proposition 10 we know that, without regulation, all firms would strictly prefer selective disclosure when they are also allowed to personalize their pricing. When consumers are wary and there is competition, regulation that bans the collection of information, and thus selective disclosure, can benefit consumers when there is not much competition, e.g., in the case where $M = 2$. Instead, with vigorous competition (large M), wary consumers are also hurt by such a ban. A strictly better policy would now be to require consumer consent. In this case, there is no longer the danger that consumer consent regulation will backfire, given that firms, with personalized pricing, strictly prefer selective disclosure, even when wary consumers observe that information collection occurs.[33] However, when consumers are unwary and when they remain unwary even after giving their

[33]Therefore, firms do not wish to commit not to engage in selective disclosure by not asking for consumer consent, contrary to what happens in the baseline case without personalized pricing (see Proposition 5).

"blanket" consent, a light-touch regulation requiring such consent does not protect unwary consumers against exploitation by a monopolist. We single this out as the only case where one could robustly argue in favor of a strict prohibition of the collection and use of personally identifiable data for the purpose of hypertargeting.

Proposition 13 (Regulation with Personalized Pricing) *Even when personalized pricing is feasible, provided that there is sufficient competition, a ban on the collection of data that allows selective disclosure will always reduce consumer surplus, irrespective of whether consumers are wary or unwary of this practice. At the opposite extreme, in a monopoly market in which requiring consent does not make consumers wary of selective disclosure, banning information collection is necessary to protect unwary consumers from being exploited.*

6 Concluding Remarks

The greater availability of personally identifiable data opens up new opportunities for firms to tailor their advertising messages to the perceived preferences of particular consumers. Thereby, marketing and selling online increasingly shares features of more traditional personalized channels, such as face-to-face interaction with a salesperson. A good salesperson in a traditional, personalized channel can use an encounter with a consumer both to learn about the preferences of the consumer and to tailor communication accordingly. When information collection happens online, are consumers sufficiently wary of hypertargeting practices? What are the implications for firms, consumers, and welfare when a previously anonymous channel becomes personalized? And what are the implications for political campaigning, where hypertargeting is increasingly used as well?

Our analysis delivers the following main insights:

- When consumers are sufficiently wary, they anticipate that firms will prominently display the attributes of a product or service that match each consumer's particular needs and taste. Wary consumers learn not only about the displayed attributes, but also, indirectly, about the attributes that firms do not display. With fixed prices, selective disclosure thus allows wary consumers to make more informed decisions, consequently increasing consumer surplus and social welfare.

- More subtly, in the absence of personalized pricing, we find that unwary consumers who remain unaware about firms' ability to selectively disclose may end up benefitting. This is the case, in particular, when an unwary consumer's inflated perception of one product is compensated for by an equal inflation in the perception of rival products.

- Given that firms may not necessarily profit from selective disclosure when it is rationally anticipated by consumers, even light-touch regulation that requires consumer consent to information collection may backfire by providing firms with a profitable commitment to abstain from collecting information, to the detriment of (wary) consumers.

- We identify one case where regulatory intervention may be warranted: when a monopolistic firm facing unwary consumers is able to both selectively communicate and price discriminate. These practices together can lead to exploitation, so that a consumer would actually be better off by choosing not to participate in this market in the first place.

- To illustrate the flexibility of our framework, we have extended it to cover political campaigning. A key difference to the application with consumers and firms is that the ultimate decision, such as the motion in a referendum or the election of a candidate, is made collectively on the basis of the aggregation of all voters' preferences. A new effect arises, because the choice of selective communication based on information about voters' preferences and political orientation affects not only the conditional expected value in case a voter is pivotal, but also the likelihood that a voter becomes pivotal in the first place. We have verified that hypertargeting by only one of the two candidates tends to reduce true voter expected welfare.

Our results are obtained in a stylized model where ex-ante symmetric firms can divulge information only about their own products or services. Our model does not apply to the case when a firm's individual disclosure allows consumers to also learn about the products of rival firms. Future research could also further untangle the two steps of collecting personally identifiable data and utilizing it to tailor advertising and pricing. The first step, in particular, may be undertaken by a third party that could then sell the information to a chosen set of firms, akin to the information broker in Taylor (2004). We also abstracted away from costs of information acquisition. When such costs are positive and non-negligible, how will the incentives of firms to collect information compare to the first-best level, depending on their ability to increase differentiation by tailoring information?

Such costs of information acquisition are the focus of the larger law and economics literature on transparency. Incentives to collect information may be too high when the prime purpose of information is to affect the distribution of surplus (Hirshleifer 1971), as is possibly the case when information allows firms to better price discriminate.[34] To better trade-off

[34]The literature on law and economics has also discussed more broadly the benefits of greater transparency for expanding efficiency-enhancing trade (Stigler 1980, Posner 1981). Hermalin and Katz (2006) show, however, that trade efficiency may not monotonically increase with information.

the social costs and benefits of collecting and using personally identifiable data, instead of prohibiting these practices, it has been proposed to essentially grant agents property rights over such information (e.g., Shapiro and Varian 1997). Our analysis reveals a particular twist to this policy. We show that a policy that requires consumer consent may allow firms to commit to abstain from selective communication even when this would benefit consumers. Finally, a different twist on the costs of transparency has been recently offered in the marketing literature on targeted advertising, which allows firms to better restrict the scope of their marketing to those consumers who are likely to purchase in the first place (cf. Athey and Gans 2010 for its impact on media competition). Several recent papers in marketing (e.g., Goldfarb and Tucker 2011; Campbell, Goldfarb, and Tucker 2011) analyze, both theoretically and empirically, how more restrictive privacy rights affect competition and welfare by potentially making advertising campaigns less cost-effective. Combined with the insights from our analysis, the protection of privacy rights should thus always be considered while taking into account competition and its benefits to consumers.

References

Abse, Nathan. 2013. *Big Data and Microtargeted Political Advertising in Election 2012: The Challenge Ahead*. Interactive Advertising Bureau, http://www.iab.net/political advertising

Acquisti, Alessandro, and Hal R. Varian. 2005. Conditioning Prices on Purchase History. *Marketing Science*, 24(3), 367–381.

An, Mark Yuying. 1998. Logconcavity versus Logconvexity: A Complete Characterization. *Journal of Economic Theory*, 80(2), 350–369.

Anderson, Simon, and André de Palma. 2012. Competition for Attention in the Information (Overload) Age. *RAND Journal of Economics*, 43(1), 1–25.

Athey, Susan, and Joshua Gans. 2010. The Impact of Targeting Technology on Advertising and Media Competition. *American Economic Review*, 100(2), 608–613.

Bagnoli, Mark, and Ted Bergstrom. 2005. Log-concave Probability and its Applications. *Economic Theory*, 26(2), 445–469.

Bagwell, Kyle. 2007. The Economic Analysis of Advertising. In: Mark Armstrong and Robert Porter (eds.), *Handbook of Industrial Organization, Volume 3*, Chapter 28, 1701–1844.

Board, Simon. 2009. Revealing Information in Auctions: The Allocation Effect. *Economic Theory*, 38(1), 125–135.

Butters, Gerard R. 1977. Equilibrium Distributions of Sales and Advertising Prices. *Review of Economic Studies*, 44(3), 465–491.

Caillaud, Bernard, and Jean Tirole. 2007. Consensus Building: How to Persuade a Group. *American Economic Review*, 97(5), 1877–1900.

Campbell, James, Goldfarb, Avi, and Catherine Tucker. 2011. Privacy Regulation and Market Structure. Mimeo.

Che, Yeon-Koo, Wouter Dessein, and Navin Kartik. 2013. Pandering to Persuade. *American Economic Review*, 103(1), 47–79.

Chakraborty, Archishman, and Rick Harbaugh. 2007. Comparative Cheap Talk. *Journal of Economic Theory*, 132(1), 70–94.

Chakraborty, Archishman, and Rick Harbaugh. 2010. Persuasion by Cheap Talk. *American Economic Review*, 100(5), 2361–2382.

Coate, Stephen. 2004. Political Competition with Campaign Contributions and Informative Advertising. *Journal of the European Economic Association*, 2(5), 772–804.

DellaVigna, Stefano, and Matthew Gentzkow. 2010. Persuasion: Empirical Evidence. *Annual Review of Economics*, 2, 643–669.

Dessein, Wouter, Andrea Galeotti, and Tano Santos. 2012. Rational Inattention and Organizational Focus. Mimeo.

Dixit, Avinash and Victor Norman. 1978. Advertising and Welfare. *Bell Journal of Economics*, 9(1), 1–17.

Eyster, Erik and Matthew Rabin. 2005. *Cursed Equilibrium*, Econometrica, 73(5), 1623–1672.

Fishman, Michael J., and Kathleen M. Hagerty. 1990. The Optimal Amount of Discretion to Allow in Disclosure. *Quarterly Journal of Economics*, 105(2), 427–444.

Federal Trade Commission. 2012. *Protecting Consumer Privacy in an Era of Rapid Change. Recommendations for Businesses and Policy Makers.* http://www.ftc.gov/os/2012/03/120326privacyreport.pdf

Galeotti, Andrea, and Andrea Mattozzi. 2011. 'Personal Influence': Social Context and Political Competition. *American Economic Journal: Microeconomics*, 3(1), 307–327.

Ganuza, Juan-José and José S. Penalva. 2010. Signal Orderings based on Dispersion and the Supply of Private Information in Auctions. *Econometrica*, 78(3), 1007-1030.

Glazer, Jacob, and Ariel Rubinstein. 2004. On Optimal Rules of Persuasion. *Econometrica*, 72(6), 1715–1736.

Goldfarb, Avi, and Catherine Tucker. 2011. Privacy Regulation and Online Advertising. *Management Science*, 57(1), 57–71.

Grossman, Gene M. and Carl Shapiro. 1984. Informative Advertising with Differentiated Products. *Review of Economic Studies*, 51(1), 63–81.

Grossman, Sanford. 1981. The Informational Role of Warranties and Private Disclosure about Product Quality. *Journal of Law and Economics*, 24(3), 461–83.

Hermalin, Benjamin, and Michael Katz. 2006. Privacy, Property Rights, and Efficiency: The Economics of Privacy as Secrecy. *Quantitative Marketing and Economics,* 4(3), 209–239.

Hirshleifer, David A., Sonya S. Lim, and Siew Hong Teoh. 2004. Disclosure to an Audience with Limited Attention. SSRN Working Paper.

Hirshleifer, Jack. 1971. The Private and Social Value of Information and the Reward to Inventive Activity. *American Economic Review*, 61(4), 561–574.

Johnson, Justin P. 2013. Targeted Advertising and Advertising Avoidance. *RAND Journal of Economics*, 44(1), 128–144.

Johnson, Justin P., and David P. Myatt. 2006. On the Simple Economics of Advertising, Marketing, and Product Design. *American Economic Review*, 96(3), 756–784.

Jovanovic, Boyan. 1982. Truthful Disclosure of Information. *Bell Journal of Economics*, 13(1), 36–44.

Kamenica, Emir, and Matthew Gentzkow. 2011. Bayesian Persuasion. *American Economic Review*, 101(6), 2590–2615.

Kamenica, Emir, Mullainathan, Sendhil and Richard Thaler. 2011. Helping Consumers Know Themselves, *American Economic Review: Papers and Proceedings*, 101(3), 417–422.

Kendall, Chad, Nannicini, Tommaso and Francesco Trebbi. 2013. *How Do Voters Respond to Information? Evidence from a Randomized Campaign*. NBER Working Paper No. 18986.

Lancaster, Kelvin. 1966. A New Approach to Consumer Theory. *Journal of Political Economy*, 74(2), 132–157.

Lewis, Tracy R., and David E. M. Sappington. 1994. Supplying Information to Facilitate Price Discrimination. *International Economic Review*, 35(2), 309–327.

Milgrom, Paul. 1981. Good News and Bad News: Representation Theorems and Applications. *Bell Journal of Economics*, 12(2), 380–391.

Milgrom, Paul, and John Roberts. 1986. Relying on Information of Interested Parties. *RAND Journal of Economics*, 17(1), 18–32.

Posner, Richard A. 1981. The Economics of Privacy. *American Economic Review*, 71(2), 405–409.

Prat, Andrea. 2006. Rational Voters and Political Advertising. In: Barry R. Weingast and Donald A. Wittman (eds.), *The Oxford Handbook of Political Economy*, Chapter 3, 50–63.

Rayo, Luis, and Ilya Segal. 2010. Optimal Information Disclosure. *Journal of Political Economy*, 118(5), 949–987.

Simon, Herbert A. 1971. Designing Organizations for an Information-Rich World. In: Martin Greenberger (ed.), *Computers, Communication, and the Public Interest*, Baltimore, MD: The Johns Hopkins Press.

Sims, Christopher A. 2003. Implications of Rational Inattention. *Journal of Monetary Economics*, 50(3), 665–690.

Shapiro, Carl and Hal R. Varian. 1997. US Government Information Policy. Unpublished. University of California, Berkeley.

Stigler, George J. 1980. An Introduction to Privacy in Economics and Politics. *Journal of Legal Studies*, 9(4), 623–644.

Taylor, Curtis. 2004. Consumer Privacy and the Market for Customer Information. *RAND Journal of Economics*, 35(4), 631–651.

Van Zandt, Timothy. 2004. Information Overload in a Network of Targeted Communication, *RAND Journal of Economics*, 35(3), 542–560.

Villas-Boas, Miguel J. 1999. Dynamic Competition with Customer Recognition. *RAND Journal of Economics*, 30(4), 604–631.

White House. 2012. *Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy.* Washington, DC, February. http://www.whitehouse.gov/sites/default/files/privacy-final.pdf

Appendix A: Omitted Proofs and Derivations

Proof of Observation 2. Let us first show that, when f is logconcave, the distribution of a wary consumer's posterior valuation with selective disclosure has more mass in the upper tail, compared to the distribution when disclosure is non-selective, also when $f(u)$ has unbounded upper support such that $\overline{U} = \infty$. Thus, we want to show that $N'(U) = n(U)$ goes to zero faster for $U \to \infty$ than $S'(U) = s(U)$, or, equivalently, that $\lim_{U\to\infty}(n(U)/s(U)) < 1$. We have

$$\frac{n(U)}{s(U)} = \frac{f(U - E[u])}{2F(u_S(U))f(u_S(U))} + \frac{f(U - E[u])u_S(U)}{2F^2(u_S(U))} - \frac{f(U - E[u])E[u \mid u \leq u_S(U)]}{2F^2(u_S(U))}, \quad (16)$$

where $u_S(U)$ solves (3). Now, as $U \to \infty$, we have that $u_S(U) \to \infty$ and, hence, $F(u_S(U)) \to 1$. Further, for $U \to \infty$, $f(U - E[u])$ goes to zero at least exponentially fast.[35] Hence, the third term in (16) goes to zero as $U \to \infty$, for the second term we have

$$\lim_{U\to\infty} \frac{f(U - E[u])u_S(U)}{2F^2(u_S(U))} = \lim_{U\to\infty} \frac{f(U - E[u])\left(U - E[u \mid u \leq u_S(U)]\right)}{2F^2(u_S(U))} = 0,$$

and the limit of the first term is given by

$$\lim_{U\to\infty} \frac{f(U - E[u])}{2F(u_S(U))f(u_S(U))} = \lim_{U\to\infty} \frac{f(u_S(U) + E[u \mid u \leq u_S(U)] - E[u])}{2F(u_S(U))f(u_S(U))} = \frac{1}{2}.$$

Next, we show that single crossing holds if u is uniformly or exponentially distributed: For the uniform distribution with support $[\underline{u}, \overline{u}]$, $N(U)$ and $S(U)$ are given by

$$N(U) = \frac{2U - 3\underline{u} - \overline{u}}{2\left(\overline{u} - \underline{u}\right)} \text{ and } S(U) = \left(\frac{2\left(U - 2\underline{u}\right)}{3\left(\overline{u} - \underline{u}\right)}\right)^2,$$

which clearly can have at most two intersections. It is easily verified that these occur at $\overline{U} = \overline{u} + E[u] = \left(3\overline{u} + \underline{u}\right)/2$ and at $\widetilde{U} = \left(3\overline{u} + 5\underline{u}\right)/4$, with $\underline{u} + E[u] < \widetilde{U} < \overline{u} + E[u]$. For the exponential distribution with parameter λ, $N(U)$ and $S(U)$ are given by

$$N(U) = 1 - e^{1-\lambda U} \text{ and } S(U) = \left(1 - e^{-\lambda u_S(U)}\right)^2,$$

where $u_S(U)$ solves

$$U = u_S + \frac{1}{\lambda} - \frac{e^{-\lambda u_S}}{1 - e^{-\lambda u_S}} u_S. \quad (17)$$

To show that $S(U)$ and $N(U)$ cross exactly once in the interior of both supports, note that we have $\Delta(U) = S(U) - N(U)$ given by

$$\Delta(U) = \exp\left\{-\lambda u_S(U)\right\}\left[\exp\left\{-\lambda u_S(U)\right\} + \exp\left\{\lambda \frac{e^{-\lambda u_S(U)} u_S(U)}{1 - e^{-\lambda u_S(U)}}\right\} - 2\right],$$

[35] It is well known that logconcave densities have at most an exponential tail, i.e., $f(u) = o(e^{-\mu u})$ for $u \to \infty$; see An (1998).

which uses (17). Clearly, as $u_S(U)$ is strictly monotonic in U, $\Delta(U) \to 0$ as $U \to \infty$. For bounded U, $\Delta(U) = 0$ if and only if the term in brackets is equal to zero, i.e., in case u_S solves

$$\exp\{-\lambda u_S\} + \exp\left\{\lambda \frac{e^{-\lambda u_S} u_S}{1 - e^{-\lambda u_S}}\right\} = 2. \tag{18}$$

Let us show that this equation has a unique solution. Taking the derivative of the left-hand side with respect to u_S gives

$$-\lambda \exp\{-\lambda u_S\} - \lambda e^{-\lambda u_S} \frac{e^{-\lambda u_S} + \lambda u_S - 1}{\left(1 - e^{-\lambda u_S}\right)^2} \exp\left\{\lambda u_S \frac{e^{-\lambda u_S}}{1 - e^{-\lambda u_S}}\right\},$$

which is clearly negative if

$$g(u_S) := e^{-\lambda u_S} + \lambda u_S \geq 1.$$

From

$$g'(u_S) = -\lambda e^{-\lambda u_S} + \lambda = \lambda\left(1 - e^{-\lambda u_S}\right),$$

$g(u_S)$ is strictly increasing for $u_S \geq 0$ and thus minimized at $u_S = 0$. Together with $g(u_S = 0) = 1$, it then follows that $g(u_S) \geq 1$ for all $u_S \geq 0$. Hence, (18) has at most one solution. But from $\Delta(U = 1/\lambda) = S(1/\lambda) > 0$ and $\Delta(U) \to 0$ for $U \to \infty$, together with the fact that $E[U]$ is the same with selective and non-selective disclosure, we must have a single crossing of $S(U)$ and $N(U)$ for $\frac{1}{\lambda} < U < \infty$. **Q.E.D.**

Proof of Proposition 2. It remains to deal with the assertions for the monopoly case. First, assume that u_i has a uniform distribution. The difference between an unwary consumer's ex-ante expected utility with non-selective disclosure and the one with selective disclosure,

$$\begin{gathered}\left[F(R - E[u]) - F^2(R - E[u])\right] R \\ + \int_{R-E[u]}^{\overline{u}} \left[u_d + E[u]\right] f(u_d) du_d - \int_{R-E[u]}^{\overline{u}} \left[u_d + E[u \mid u \leq u_d]\right] 2 f(u_d) F(u_d) du_d,\end{gathered} \tag{19}$$

is then exactly zero, once we substitute for the distribution function. For the exponential distribution, after some transformations expression (19) becomes

$$e^{-\lambda\left(R - \frac{1}{\lambda}\right)}\left[R e^{-\lambda\left(R - \frac{1}{\lambda}\right)} - \frac{1}{\lambda}\right].$$

Now note that $Re^{-\lambda\left(R - \frac{1}{\lambda}\right)}$, with $R \geq \frac{1}{\lambda}$, is maximized at $R = \frac{1}{\lambda} = E[u]$, where it takes on the value $\frac{1}{\lambda}$. From this we see that an unwary customer is always (weakly) better off under selective disclosure, and indifferent only if he purchases always ($R \leq E[u] = \frac{1}{\lambda}$) or never ($R \to \infty$). **Q.E.D.**

Proof of Proposition 7. It remains to show the results for the uniform distribution. First, we will show that candidates prefer selective disclosure, as this maximizes their probability

of getting elected (Q). Note first that Q is maximized by maximizing q, the probability of a single voter electing the respective candidate. If both candidates choose the same communication strategy, we clearly have $q = \frac{1}{2}$. If, however, one candidate ($m = 1$) discloses selectively and the other one ($m = 2$) non-selectively, any voter will elect $m = 1$ with probability

$$q_S = \int_{\underline{U}}^{\overline{U}} (1 - S(U))\, dN(U) = \int_{\underline{u}+E[u]}^{\overline{u}+E[u]} S'(U)\, N(U) dU,$$

where we have used integration by parts. Substituting for $N(U)$ and $S'(U)$ from (6) gives, after some algebra, $q_S = \frac{14}{27} > \frac{1}{2}$. Next, note that when both candidates choose non-selective disclosure, we have $y = y_N$ as given in (12) (with $q_N = \frac{1}{2}$) and

$$E_N[U^{(1)}] = \int_{\underline{u}+E[u]}^{\overline{u}+E[u]} U d(N)^2(U),$$

which for the uniform distribution becomes ultimately

$$E_N[U^{(1)}] = \frac{7\overline{u} + 5\underline{u}}{6}.$$

If only $m = 1$ chooses selective disclosure, $y = y_S$ is obtained from (12) with $q_S = \frac{14}{27}$ and we have

$$\begin{aligned} E_S[U^{(1)}] &= \int_{\underline{u}+E[u]}^{\overline{u}+E[u]} \left[\int_{2\underline{u}}^{\overline{u}+E[u]} \max\left\{U^1, U^2\right\} dS(U^1) \right] dN(U^2) \\ &= 2S(\underline{u} + E[u])E[u] + \int_{\underline{u}+E[u]}^{\overline{u}+E[u]} \left[\int_{\underline{u}+E[u]}^{U^2} U^2 dS(U^1) + \int_{U^2}^{\overline{u}+E[u]} U^1 dS(U^1) \right] dN(U^2), \end{aligned}$$

which in the uniform case yields ultimately

$$E_S[U^{(1)}] = \frac{32\overline{u} + 22\underline{u}}{27}.$$

Substituting into condition (14), this becomes, again after several transformations, $\left(\frac{729}{728}\right)^{\frac{V-1}{2}} < \frac{30}{27}$ and yields the cutoff $\widetilde{V} = 153$. **Q.E.D.**

Proof of Proposition 8. We first show that, when voters are unwary, candidates always want to hypertarget. To see this, assume that only one candidate ($m = 1$) hypertargets. Then, any voter will elect $m = 1$ with probability

$$\begin{aligned} \widehat{q} &= \int_{\underline{U}}^{\overline{U}} (1 - S_n(U))\, dN(U) = \int_{\underline{u}+E[u]}^{\overline{u}+E[u]} S_n'(U)\, N(U) dU \\ &= \int_{\underline{u}+E[u]}^{\overline{u}+E[u]} 2f(U - E[u])F^2(U - E[u]) dU = \left[\frac{2}{3}F^3(U - E[u])\right]_{\underline{u}+E[u]}^{\overline{u}+E[u]} = \frac{2}{3}, \end{aligned}$$

which is strictly larger than the corresponding probability when both candidates employ the same communication strategy.

Next, the assertion that unwary voters are best off when both candidates hypertarget follows from Proposition 7, after recognizing that a wary customer's expected utility provides an upper bound for a naive customer's expected utility for any given communication strategy. The result then follows because both the probability of becoming pivotal (y), as well as the ex-ante expected utility conditional on being pivotal ($E[U^{(1)}]$), are the same for unwary and wary voters, if both candidates chose the same communication strategy. Finally, consider the remaining comparison of the case where communication strategies are asymmetric and the one where both candidates communicate non-selectively. Assume that $m = 2$ always communicates non-selectively. Then the difference between a pivotal voter's ex-ante expected utility when $m = 1$ also communicate non-selectively and when $m = 1$ communicates selectively is equal to

$$\int_{\underline{U}}^{\overline{U}} \left[\begin{array}{c} \left[F(U^2 - E[u]) - F^2(U^2 - E[u])\right] U^2 \\ + \int_{U^2-E[u]}^{\overline{u}} \left[u_d^1 + E[u]\right] dF(u_d^1) - \int_{U^2-E[u]}^{\overline{u}} \left[u_d^1 + E[u \mid u \leq u_d^1]\right] dF^2(u_d^1) \end{array} \right] dN(U^2),$$

where the integrand becomes exactly zero once we substitute for the uniform distribution function (cf. also the proof of Proposition 2). Hence, in case preferences are uniformly distributed, $E[U^{(1)}]$ is unaffected by the disclosure strategy of $m = 1$. However, as the probability of becoming pivotal (y) is smaller in the asymmetric case the result follows. **Q.E.D.**

Proof of Proposition 9. Note first that the expected utility of a pivotal voter, no matter whether he is wary or unwary, is independent of the number of unwary voters V_u. Hence, from (13), V_u can only affect expected voter welfare through its effect on the probability of becoming pivotal. If both candidates choose the same communication strategy, both unwary and wary voters will elect each candidate with the same ex-ante probability. Hence, the probability of becoming pivotal is given by (12) with $q = \frac{1}{2}$ and, thus, only depends on the total number of voters (and not on V_u). When, however, only one candidate, say $m = 1$, discloses selectively, then by using the function $u_S(U)$ we can generally derive the following bound for the probability with which a wary voter elects $m = 1$:

$$q = \int_{\underline{U}}^{\overline{U}} \left[1 - S\left(U\right)\right] dN\left(U\right) \leq \left[\frac{2}{3} F^3(u_S)\right]_{\underline{u}+E[u]}^{\overline{u}+E[u]} < \frac{2}{3} = \widehat{q}.$$

When now $q \geq 1/2$, this implies that the likelihood of becoming pivotal decreases with V_u, as is intuitive. To see this formally, pick an arbitrary voter v' and determine how the probability of the remaining $V - 1$ voters generating a draw changes when a single voter $v \neq v'$ is wary

versus when v is unwary.[36] To do so, decompose the probability of a draw among these $V-1$ voters according to the voting decision of voter v; denote the number of votes out of $V\backslash\{v,v'\}$ cast for candidate $m=1$ by V_1 and the number of votes cast for $m=2$ by V_2. Then, if v votes for $m=1$, the probability of a draw is given by $X=\Pr\left(V_1=\frac{V-1}{2}-1, V_2=\frac{V-1}{2}\right)$. If v votes for $m=2$, the respective probability is given by $Y=\Pr\left(V_1=\frac{V-1}{2}, V_2=\frac{V-1}{2}-1\right)$. Denoting the probability with which voter v votes for $m=1$ by q_v, the total probability of a draw among these $V-1$ voters is thus $q_vX+(1-q_v)Y$. Now note that from $q\geq\frac{1}{2}$ and $\widehat{q}>\frac{1}{2}$ it follows that $X<Y$,[37] implying that $q_vX+(1-q_v)Y$ is larger for $q_v=q$ than for $q_v=\widehat{q}>q$. **Q.E.D.**

Proof of Proposition 11. The result for $M=1$ is immediate. Consider the case with $M=2$ and pick one firm m with U^m distributed according to $G(\cdot)$. Given the realization U^n, $n\neq m$, after integration by parts a customer's expected utility is

$$\left[1-G(U^n)\right]U^n+\int_{\underline{U}}^{U^n}UdG(U)=U^n-\int_{\underline{U}}^{U^n}G(U)dU.$$

Denoting the distribution of U^n by $G^n(\cdot)$, the difference in a customer's ex-ante expected utility when firm m communicates selectively rather than non-selectively is equal to

$$\int_{\underline{U}}^{\overline{U}}\left[\int_{\underline{U}}^{U^n}\left[N(U)-S(U)\right]dU\right]dG^n(U^n).$$

From the single-crossing property of $N(U)$ and $S(U)$, the integral in brackets is strictly negative for all interior U^n, so that the total expression is strictly negative. Thus, we conclude that, with $M=2$, a wary customer is worse off under selective disclosure by firm m, independent of what the competing firm n does.

Next, consider the case with sufficiently high values of M. Proceeding as before, pick any firm m with U^m distributed according to $G(\cdot)$. Recall that we denote the maximum of expected utilities over all other firms by $U^{(1:M\backslash m)}$ and, likewise, the second-highest realization by $U^{(2:M\backslash m)}$, with corresponding distributions $G^{(1:M\backslash m)}(\cdot)$ and $G^{(2:M\backslash m)}(\cdot)$. Given that a customer's expected utility with firm m is independent of the realizations of all other $M-1$ firms, for any two realizations $U^{(1:M\backslash m)}$ and $U^{(2:M\backslash m)}$, a customer's expected utility is

$$G(U^{(2:M\backslash m)})U^{(2:M\backslash m)}+\left[1-G(U^{(1:M\backslash m)})\right]U^{(1:M\backslash m)}+\int_{U^{(2:M\backslash m)}}^{U^{(1:M\backslash m)}}UdG(U),$$

[36] For $v=v'$ it is trivial that the probability of v' becoming pivotal is unaffected.

[37] Note that for each possible outcome with $V_1=(V-1)/2-1$ and $V_2=(V-1)/2$ there exists a respective outcome with $V_1=(V-1)/2$ and $V_2=(V-1)/2-1$ where all but one voter (call this voter v'') take the same election decision. But as v'' is more likely to vote for $m=1$, the result follows.

which can be transformed to obtain

$$U^{(1:M\setminus m)} - \int_{U^{(2:M\setminus m)}}^{U^{(1:M\setminus m)}} G(U)dU.$$

Now, denoting the joint distribution of $U^{(1:M\setminus m)}$ and $U^{(2:M\setminus m)}$ by $H(\cdot,\cdot)$, the difference in a customer's ex-ante expected utility when firm m communicates selectively rather than non-selectively is

$$\int_{U^{(1:M\setminus m)}} \int_{U^{(2:M\setminus m)}} \left[\int_{U^{(2:M\setminus m)}}^{U^{(1:M\setminus m)}} [N(U) - S(U)]\, dU \right] dH(U^{(1:M\setminus m)}, U^{(2:M\setminus m)}), \tag{20}$$

where the term in brackets can be rewritten as

$$\int_{U^{(2:M\setminus m)}}^{U^{(1:M\setminus m)}} z(U)dU = -\int_{\underline{U}}^{U^{(2:M\setminus m)}} z(U)dU - \int_{U^{(1:M\setminus m)}}^{\overline{U}} z(U)dU,$$

with $z(U) := N(U) - S(U)$. Integrating by parts, (20) becomes

$$\int_{\underline{U}}^{\overline{U}} z(U) \left[G^{(2:M\setminus m)}(U) - G^{(1:M\setminus m)}(U) \right] dU, \tag{21}$$

which we will show to be positive for M large.

To see this, let us first introduce some notation. Denote the number of firms among the other $M-1$ firms that communicate selectively by M_S and the number of firms communicating non-selectively by M_N, with $M_S + M_N = M - 1$. Then, clearly, we have

$$G^{(1:M\setminus m)}(U) = S^{M_S}(U) N^{M_N}(U).$$

Note next that the distributions of the highest and second-highest value from all M_S firms communicating selectively are given by

$$\begin{aligned} S^{(1)}(U) &= S^{M_S}(U), \\ S^{(2)}(U) &= M_S S^{M_S-1}(U) - (M_S - 1)\, S^{M_S}(U), \end{aligned}$$

so that

$$\begin{aligned} & G^{(2:M\setminus m)}(U) \\ = \; & S^{M_S-1}(U) N^{M_N-1}(U) \left[M_S N(U) - (M_S + M_N - 1)\, S(U)N(U) + M_N S(U) \right]. \end{aligned}$$

Hence, we can write

$$\begin{aligned} & G^{(2:M\setminus m)}(U) - G^{(1:M\setminus m)}(U) \\ = \; & S^{M_S-1}(U) N^{M_N-1}(U) \left[M_S N(U) - (M_S + M_N)\, S(U)N(U) + M_N S(U) \right]. \end{aligned}$$

Substituting in (21) we can decompose this as follows:

$$\int_{\underline{U}}^{\widetilde{U}} z(U)S^{M_S-1}(U)N^{M_N-1}(U)\left[M_S N(U)+M_N S(U)-(M_S+M_N)\,S(U)N(U)\right]dU$$
$$+\int_{\widetilde{U}}^{\overline{U}} z(U)S^{M_S-1}(U)N^{M_N-1}(U)\left[M_S N(U)+M_N S(U)-(M_S+M_N)\,S(U)N(U)\right]dU,$$

where the first integral is negative and the second is positive. Now, note that from the single-crossing property there exist $U^*, U^{**} \in \left(\widetilde{U}, 1\right)$ such that $z(U) > 0$ for all $U \in [U^*, U^{**}]$. Next, defining

$$\begin{aligned} x_1 &= \min_{\left[\underline{U},\widetilde{U}\right]} \left\{z(U)\left[M_S N(U)+M_N S(U)-(M_S+M_N)\,S(U)N(U)\right]\right\} < 0, \\ x_2 &= \min_{[U^*,U^{**}]} \left\{z(U)\left[M_S N(U)+M_N S(U)-(M_S+M_N)\,S(U)N(U)\right]\right\} > 0, \end{aligned}$$

we obtain that (21) is bounded from below by

$$\begin{aligned} & x_1 \int_{\underline{U}}^{\widetilde{U}} S^{M_S-1}(U)N^{M_N-1}(U)dU + x_2 \int_{U^*}^{U^{**}} S^{M_S-1}(U)N^{M_N-1}(U)dU \\ \geq\ & x_1 \left(\widetilde{U}-\underline{U}\right) S^{M_S-1}(\widetilde{U})N^{M_N-1}(\widetilde{U}) + x_2\left(U^{**}-U^*\right) S^{M_S-1}(U^*)N^{M_N-1}(U^*) \\ =\ & \left[\frac{x_1\left(\widetilde{U}-\underline{U}\right)}{x_2\left(U^{**}-U^*\right)}\left(\frac{N(\widetilde{U})}{N(U^*)}\right)^{M_N-1}\left(\frac{S(\widetilde{U})}{S(U^*)}\right)^{M_S-1}+1\right] \\ & \cdot x_2\left(U^{**}-U^*\right) S^{M_S-1}(U^*)N^{M_N-1}(U^*). \end{aligned}$$

This is now strictly positive when either M_N or M_S become sufficiently large (or both), given that then the term in brackets is strictly positive (as the first, negative term goes to zero, so that the whole term converges to one). **Q.E.D.**

For Online Publication

Supplementary Appendix B: Different Weights for Different Attributes

In this Appendix, we introduce asymmetry in the importance of the two attributes by stipulating the following specification of utility

$$u = w_1 u_1 + w_2 u_2,$$

with $w_1 \geq w_2$. We focus on the tractable case where u_1 and u_2 are independent and uniformly distributed on $[\underline{u}, \overline{u}]$.

Selective Disclosure to Wary Consumers. For selective disclosure we restrict attention to the characterization of a rational expectations equilibrium where the disclosure rule of a firm is linear.

Lemma 1 *With selective disclosure there is a rational expectations equilibrium in which a firm follows the following linear disclosure rule: The firm discloses $d = 1$ when*

$$u_1 \geq \frac{(w_1 - w_2)}{w_1}\underline{u} + \frac{w_2}{w_1}u_2, \tag{22}$$

while otherwise it discloses $d = 2$.

Proof. We look for a linear decision rule such that the disclosed attribute is $d = 1$ whenever $u_1 \geq a + bu_2$. Note that the firm wants to maximize a wary consumer's expected valuation. If this rule is rationally anticipated, then choosing $d = 1$ is indeed optimal if and only if

$$w_1 u_1 + w_2 E\left[u_2 | u_2 \leq \frac{u_1 - a}{b}\right] \geq w_2 u_2 + w_1 E\left[u_1 | u_1 \leq a + bu_2\right],$$

which transforms to the requirement that

$$u_1 \geq \frac{\left(\frac{w_1}{2} + \frac{w_2}{2b}\right) a + (w_1 - w_2)\frac{\underline{u}}{2}}{\left(w_1 + \frac{w_2}{2b}\right)} + \frac{\left(w_2 + \frac{w_1 b}{2}\right)}{\left(w_1 + \frac{w_2}{2b}\right)} u_2.$$

Comparing coefficients yields (22). **Q.E.D.**

If rule (22) is followed and when $d = 1$ is chosen, the expected utility equals

$$U = \frac{3}{2} w_1 u_1 - \frac{1}{2}(w_1 - 2w_2)\underline{u},$$

so that $U \in \left[(w_1+w_2)\,\underline{u}, \frac{3}{2}w_1\overline{u} - \frac{1}{2}(w_1-2w_2)\,\underline{u}\right]$. With $d=2$ we obtain

$$U = \frac{3}{2}w_2u_2 + \frac{1}{2}(2w_1-w_2)\,\underline{u},$$

so that now $U \in \left[(w_1+w_2)\,\underline{u}, \frac{3}{2}w_2\overline{u} + \frac{1}{2}(2w_1-w_2)\,\underline{u}\right]$. Note that from $w_1 \geq w_2$, which we stipulated without loss of generality, the highest value of U is attained when disclosing $u_1 = \overline{u}$:

$$\overline{U} = \frac{3}{2}w_1\overline{u} - \frac{1}{2}(w_1-2w_2)\,\underline{u},$$

while likewise

$$\underline{U} = (w_1+w_2)\,\underline{u}.$$

We next derive the distribution of U over this interval, restricting attention to the asymmetric case with $w_1 > w_2$.

Lemma 2 *With selective disclosure where a firm follows a linear decision rule, the ex-ante distribution of a wary consumer's expected utility is*

$$S(U) = \begin{cases} \frac{1}{w_1w_2}\left(\frac{2(U-(w_1+w_2)\underline{u})}{3(\overline{u}-\underline{u})}\right)^2 & \text{for } \underline{U} \leq U \leq U' \\ \frac{1}{w_1}\frac{2(U-(w_1+w_2)\underline{u})}{3(\overline{u}-\underline{u})} & \text{for } U' < U \leq \overline{U} \end{cases} \tag{23}$$

where

$$U' = \frac{3}{2}w_2\overline{u} + \frac{1}{2}(2w_1-w_2)\,\underline{u}, \tag{24}$$

with $U' \in \left(\underline{U}, \overline{U}\right)$ *for* $w_1 > w_2$.

Proof. From the preceding observations we obtain that the expected utilities $U \in [\underline{U}, U']$ with U' defined in (24) are realized in two possible ways: when either a respective value u_1 for $d=1$ or u_2 for $d=2$ is disclosed. Instead, the values $U \in \left(U', \overline{U}\right]$ can only be attained by disclosing a sufficiently high u_1. These observations are now used to piecewise derive the ex-ante distribution for U. Here, for $U \in \left(U', \overline{U}\right]$, we trivially have

$$S(U) = F(u_1(U)),$$

where

$$u_1(U) = \frac{2}{3w_1}U + \frac{w_1-2w_2}{3w_1}\underline{u}.$$

Suppose next $U \in [\underline{U}, U']$. A particular value of U in this interval can be attained by either choosing $d=1$ and disclosing $u_1(U)$, which is optimal, whenever

$$u_2 \leq \frac{w_1}{w_2}u_1(U) - \frac{(w_1-w_2)}{w_2}\underline{u} = \frac{2}{3w_2}U - \frac{2w_1-w_2}{3w_2}\underline{u},$$

or, by choosing $d = 2$ and disclosing

$$u_2(U) = \frac{2}{3w_2}U - \frac{2w_1 - w_2}{3w_2}\underline{u}$$

whenever

$$u_1 \leq \frac{w_1 - w_2}{w_1}\underline{u} + \frac{w_2}{w_1}u_2(U) = \frac{2}{3w_1}U + \frac{w_1 - 2w_2}{3w_1}\underline{u}.$$

So, we obtain for the distribution function on $[\underline{U}, U']$ that

$$\begin{aligned} S(U) &= \int_{\underline{U}}^{U} \left[\frac{1}{(\overline{u} - \underline{u})^2} \left(\frac{2}{3w_2}Y - \frac{2w_1 - w_2}{3w_2}\underline{u} - \underline{u} \right) \frac{2}{3w_1} \right] dY \\ &\quad + \int_{\underline{U}}^{U} \left[\frac{1}{(\overline{u} - \underline{u})^2} \left(\frac{2}{3w_1}Y + \frac{w_1 - 2w_2}{3w_1}\underline{u} - \underline{u} \right) \frac{2}{3w_2} \right] dY \\ &= \frac{1}{w_1 w_2} \left(\frac{2\left(U - (w_1 + w_2)\,\underline{u}\right)}{3\left(\overline{u} - \underline{u}\right)} \right)^2. \end{aligned}$$

Q.E.D.

Non-selective Disclosure. Note that with non-selective disclosure a firm applies the same disclosure rule to each consumer. We now determine what is the optimal disclosure rule in this case. For this we must specify whether there is competition or not. We focus here on the case of monopoly.

If, say, the rule is to always disclose the first attribute $d = 1$, the probability of a purchase is given by

$$\Pr\left[w_1 u_1 + w_2 E[u_2] \geq R\right] = 1 - F\left(\frac{R - w_2 \frac{(\overline{u} + \underline{u})}{2}}{w_1} \right),$$

while, following disclosure of u_2, the purchase probability is given by

$$\Pr\left[w_2 u_2 + w_1 E[u_1] \geq R\right] = 1 - F\left(\frac{R - w_1 \frac{(\overline{u} + \underline{u})}{2}}{w_2} \right).$$

So, which attribute is optimally disclosed depends on the outside option, with disclosure of u_1 (the more important attribute) being optimal when

$$R \geq (w_1 + w_2)\,\frac{(\overline{u} + \underline{u})}{2}, \tag{25}$$

i.e., when there is no purchase without information.[38]

[38] To ensure that the consumer will both purchase and not purchase with strictly positive probability we impose the restriction that $w_2\underline{u} + w_1\frac{(\overline{u}+\underline{u})}{2} < R < w_2\overline{u} + w_1\frac{(\overline{u}+\underline{u})}{2}$.

Lemma 3 *When there is non-selective disclosure, the ex-ante distribution of a consumer's expected utility is obtained as follows:*
(i) When (25) holds, the monopolistic firm chooses $d = 1$ and we have

$$N(U) = \frac{2U - w_2(\overline{u} + \underline{u}) - 2w_1\underline{u}}{2w_1(\overline{u} - \underline{u})} \tag{26}$$

for $U \in \left[w_1\underline{u} + w_2\frac{(\overline{u}+\underline{u})}{2}, w_1\overline{u} + w_2\frac{(\overline{u}+\underline{u})}{2}\right]$.
(ii) When (25) does not hold, the firm chooses $d = 2$ and we have

$$N(U) = \frac{2U - w_1(\overline{u} + \underline{u}) - 2w_2\underline{u}}{2w_2(\overline{u} - \underline{u})} \tag{27}$$

for $U \in \left[w_2\underline{u} + w_1\frac{(\overline{u}+\underline{u})}{2}, w_2\overline{u} + w_1\frac{(\overline{u}+\underline{u})}{2}\right]$.

Selective Disclosure to Unwary Consumers. Facing an unwary customer, it is optimal to choose $d = 1$, so as to maximize the perceived valuation, when

$$w_1u_1 + w_2E[u_2] \geq w_1E[u_1] + w_2u_2,$$

which transforms to

$$u_1 \geq \frac{(w_1 - w_2)}{w_1}\frac{(\overline{u} + \underline{u})}{2} + \frac{w_2}{w_1}u_2,$$

and otherwise to disclose $d = 2$.

Lemma 4 *With selective disclosure, the ex-ante distribution of an unwary consumer's perceived expected utility $\widehat{U}$ is given by*

$$S_n(\widehat{U}) = \begin{cases} \frac{\left[2\widehat{U} - (w_1+w_2)\left(\frac{(\overline{u}+\underline{u})}{2} + \underline{u}\right)\right]^2 - \left[(w_1-w_2)\left(\frac{(\overline{u}+\underline{u})}{2} - \underline{u}\right)\right]^2}{4w_1w_2(\overline{u}-\underline{u})^2} & \text{for } \underline{\widehat{U}} \leq U \leq \widehat{U}' \\ \frac{2\widehat{U} - w_2(\overline{u}+\underline{u}) - 2w_1\underline{u}}{2w_1(\overline{u}-\underline{u})} & \text{for } \widehat{U}' < U \leq \overline{\widehat{U}} \end{cases} \tag{28}$$

where

$$\underline{\widehat{U}} = w_2\underline{u} + w_1\frac{(\overline{u} + \underline{u})}{2}, \tag{29}$$

$$\overline{\widehat{U}} = w_1\overline{u} + w_2\frac{(\overline{u} + \underline{u})}{2}, \tag{30}$$

$$\widehat{U}' = w_2\overline{u} + w_1\frac{(\overline{u} + \underline{u})}{2}, \tag{31}$$

with $\widehat{U}' \in \left(\underline{\widehat{U}}, \overline{\widehat{U}}\right)$ *for* $w_1 > w_2$.

Proof. From the preceding observations we have that, when u_1 is disclosed, the perceived utility is given by

$$\widehat{U} = w_1 u_1 + w_2 E[u_2] = w_1 u_1 + w_2 \frac{(\overline{u} + \underline{u})}{2},$$

while, when u_2 is disclosed, we obtain

$$\widehat{U} = w_2 u_2 + w_1 E[u_1] = w_2 u_2 + w_1 \frac{(\overline{u} + \underline{u})}{2}.$$

Comparing the respective bounds for $w_1 > w_2$, we obtain that values $\widehat{U} > \widehat{U}'$, where $\widehat{U}'$ is defined in (31), can only be attained by disclosing a high enough value of u_1. So, the highest value of $\widehat{U}$ is attained when disclosing $u_1 = \overline{u}$, and is given by (30). Further, observe that disclosing $u_2 = \underline{u}$ dominates disclosing any $u_1 < \frac{w_2}{w_1}\underline{u} + \frac{(w_1 - w_2)}{w_1}\frac{(\overline{u}+\underline{u})}{2}$, as it generates a higher $\widehat{U}$. Hence, the lowest value of $\widehat{U}$ is attained when disclosing $u_2 = \underline{u}$ and is given by (29). Values $\widehat{U} \in \left[\underline{\widehat{U}}, \widehat{U}'\right]$ can be realized both by disclosing a particular value of u_1 or a particular value of u_2. We thus obtain in this case for the respective distribution function

$$\begin{aligned} S_n(\widehat{U}) &= \int_{\underline{\widehat{U}}}^{\widehat{U}} \left[\frac{1}{(\overline{u} - \underline{u})^2} \left(\frac{1}{w_2} Y - \frac{w_1}{w_2} \frac{(\overline{u} + \underline{u})}{2} - \underline{u} \right) \frac{1}{w_1} \right] dY \\ &\quad + \int_{\underline{\widehat{U}}}^{\widehat{U}} \left[\frac{1}{(\overline{u} - \underline{u})^2} \left(\frac{1}{w_1} Y - \frac{w_2}{w_1} \frac{(\overline{u} + \underline{u})}{2} - \underline{u} \right) \frac{1}{w_2} \right] dY \\ &= \frac{1}{w_1 w_2} \frac{1}{4(\overline{u} - \underline{u})^2} \left[\begin{array}{c} \left(2\widehat{U} - (w_1 + w_2) \left(\frac{(\overline{u}+\underline{u})}{2} + \underline{u} \right) \right)^2 \\ - \left((w_1 - w_2) \left(\frac{(\overline{u}+\underline{u})}{2} - \underline{u} \right) \right)^2 \end{array} \right]. \end{aligned}$$

Values $\widehat{U} \in \left(\widehat{U}', \overline{\widehat{U}}\right]$ can only be realized by disclosing a sufficiently high value of u_1 and we obtain on this interval that

$$S_n(\widehat{U}) = \frac{2\widehat{U} - w_2(\overline{u} + \underline{u}) - 2w_1\underline{u}}{2w_1(\overline{u} - \underline{u})}.$$

Q.E.D.

Comparing Selective and Non-Selective Disclosure. We finally compare the outcomes with selective and non-selective disclosure, thereby confirming the robustness of our results with equal weights only.

Proposition 14 *Comparing the cases with and without selective disclosure, we have the following results:*
(i) When consumers are unwary, the distribution $S_n(\widehat{U})$ under selective disclosure dominates the distribution $N(U)$ under non-selective disclosure in the sense of FOSD.

(ii) When consumers are wary, the distribution $S(U)$ under selective disclosure and the distribution $N(U)$ under non-selective disclosure compare as follows: There exists a value $\widetilde{U}$ in the interior of both supports such that $S(U) > N(U)$ for $U < \widetilde{U}$ and $S(U) \leq N(U)$ for $U > \widetilde{U}$.[39]

Proof. The case with $w_1 = w_2$ is treated in the main text, so, we restrict attention to $w_1 > w_2$. We start with assertion (i) for unwary customers. Assume, first, that (25) holds. Then, comparing the support of $N(U)$ and $S_n(U)$ as given by (26) and (28), it is immediate that $S_n(U) = N(U)$ for $U \geq \widehat{U}'$, while the lower bound of the support of $N(U)$ is strictly smaller than the respective bound for $S_n(U)$. Now note that, for $U < \widehat{U}'$, $N(U)$ is linear while $S_n(U)$ is convex in the interior of the respective support, which, together with $S_n(\widehat{U}') = N(\widehat{U}')$ and the result for the lower tail, implies that $N(U) \geq S_n(U)$ for all U.

Second, assume that (25) does not hold. Then, comparing the expressions in (27) and (28) for $U \geq \widehat{U}'$, we find from $N(\widehat{U}') = 1 > S_n(\widehat{U}')$ that $N(U) \geq S_n(U)$ for all $U \geq \widehat{U}'$. Finally, note that, in this case, both distributions have the same lower support at $\underline{\widehat{U}}$. Thus, for $U < \widehat{U}'$, it follows from $N(\widehat{U}') > S_n(\widehat{U}')$, together with the fact that $N(U)$ is linear, while $S_n(U)$ is convex for $\underline{\widehat{U}} \leq U < \widehat{U}'$, that we must have $N(U) \geq S_n(U)$ also for all $U < \widehat{U}'$.

Next, we prove assertion (ii) for wary customers. Again, assume first that (25) holds. Then, comparing the support of $N(U)$ and $S(U)$ as given by (26) and (23), we have that $S(U)$ has more mass in the lower tail ($\underline{U} < w_1\underline{u} + w_2\frac{(\overline{u}+\underline{u})}{2}$) and upper tail ($\overline{U} > w_1\overline{u} + w_2\frac{(\overline{u}+\underline{u})}{2}$). For the comparison in the interior of both supports, note that $N(U') = w_2/w_1 = S(U')$, which, together with the fact that $S'(U) < N'(U)$ for $U > U'$, implies that $S(U) < N(U)$ for $U \in \left(U', \overline{U}\right)$. Now, as for $U < U'$ we have that $S(U)$ is convex, while $N(U)$ is linear, it follows from $N(U') = S(U')$ that there can be at most one intersection. Existence of a unique intersection at $\widetilde{U} \in (\underline{U}, U')$ then follows from $S'(U') > N'(U')$, together with the fact that the lower bound of the support is smaller under selective disclosure.

To complete the proof of assertion (ii), take next the case where (25) does not hold. Then, comparing the support of $N(U)$ and $S(U)$ as given by (27) and (23), we have that, still, $S(U)$ has more mass in the lower tail ($\underline{U} < w_2\underline{u} + w_1\frac{(\overline{u}+\underline{u})}{2}$) and upper tail ($\overline{U} > w_2\overline{u} + w_1\frac{(\overline{u}+\underline{u})}{2}$), implying that there must exist at least one intersection. For the comparison in the interior of both supports, we now distinguish two cases, depending on whether $N(U')$ is smaller or larger than $S(U')$. Take first the case with $N(U') \leq S(U')$, which from (27) and (23) holds whenever $w_2 \leq w_1/2$. In this case it can be shown that $S(U)$ is everywhere flatter than

[39] Note that this single-crossing condition is slightly weaker than the one in the main text (cf. Observation 2), where we required $S(U) < N(U)$ for $U > \widetilde{U}$. However, also in this example with $w_1 > w_2$, the requirement that $S(U) \leq N(U)$ for $U > \widetilde{U}$ holds strictly, except at a unique point U' when (25) holds.

$N(U)$, implying a unique intersection. To see this, note that $N(U)$ is linear everywhere with slope $1/\left(w_2\left(\overline{u}-\underline{u}\right)\right)$, while $S(U)$ is linear with slope $2/\left(3w_1\left(\overline{u}-\underline{u}\right)\right)$ for $U > U'$ and convex for $U \leq U'$ with $S'(U') = 4/\left(3w_1\left(\overline{u}-\underline{u}\right)\right)$.[40] Hence, the result follows because we are in the case with $w_2 \leq w_1/2$. Finally, take the case with $N(U') > S(U')$, which holds whenever $w_2 > w_1/2$, and recall that $N(U)$ is steeper than $S(U)$ for $U > U'$. Hence, any intersection must occur for some $U < U'$. Further, as $N(U)$ is linear and $S(U)$ is convex for $U < U'$, there can be at most two intersections. From $N(U') > S(U')$, and from the fact that the lower bound of the support is smaller under selective disclosure, it follows that the intersection must be unique. **Q.E.D.**

[40]Note that for $w_1 > 3w_2$, we have $U' < w_2\underline{u} + w_1\frac{(\overline{u}+\underline{u})}{2}$ such that we only have to compare the linear parts.

For Online Publication

Supplementary Appendix C: Disentangled Locations

This Appendix extends the primitives of the model by introducing separately the location of the attributes of a given product and the location of the attributes a given consumer would prefer.

To this end, consider one firm and suppose that two attributes (characteristics) are given by x_1, x_2 and are, from an ex-ante perspective, distributed uniformly on a Salop circle of circumference two. Each consumer has a preferred location for each attribute. The preferred location of a mass one of consumers is distributed uniformly around each circle and denoted by y_1, y_2. A particular consumer's true utility is then

$$2\alpha - \sum_{i=1}^{2} |y_i - x_i| = 2\alpha - \sum_{i=1}^{2} d_i,$$

where $\alpha > 1$ and $d_i := |y_i - x_i|$ denotes the discrepancy between characteristic x_i and the consumer's preferences y_i. With d_i, thus, distributed uniformly on $[0, 1]$, we obtain for the respective distribution of $u_i = \alpha - d_i$ the following:

$$\begin{aligned} F(u_i) &= \Pr(\alpha - d_i \leq u_i) = \Pr(d_i \geq \alpha - u_i) \\ &= u_i - (\alpha - 1), \end{aligned}$$

which is the distribution function of a uniform distribution on $[\underline{u}, \overline{u}]$ with $\underline{u} = \alpha - 1$ and $\overline{u} = \alpha$. Accordingly, this model with distances can be analyzed using the same methods as in the main text. Alternatively, for completeness, we can derive the distribution of U directly. So, assume without loss of generality that $d_1 \leq d_2$, so that, under selective disclosure, the firm reveals d_1. Then, given d_1, a wary consumer's expected valuation is

$$\begin{aligned} U &= 2\alpha - (d_1 + E[d_2 \mid d_2 \geq d_1]) \\ &= 2\alpha - \frac{1}{2}(3d_1 + 1), \end{aligned}$$

with ex-ante distribution

$$\begin{aligned} S(U) &= \Pr\left(2\alpha - \frac{1}{2}(3d_1 + 1) \leq U\right) \\ &= \left(\frac{2U + 4(1 - \alpha)}{3}\right)^2, \end{aligned}$$

for $U \in \left[2(\alpha - 1), 2\alpha - \frac{1}{2}\right]$. Next, an unwary customer's perceived expected valuation when facing selective disclosure is given by

$$\widehat{U} = 2\alpha - d_1 - E[d_2] = 2\alpha - d_1 - \frac{1}{2},$$

with ex-ante distribution

$$\begin{aligned} S_n(\widehat{U}) &= \Pr\left(2\alpha - d_1 - \frac{1}{2} \leq \widehat{U}\right) \\ &= \left(\widehat{U} + \frac{3}{2} - 2\alpha\right)^2, \end{aligned}$$

for $\widehat{U} \in \left[2\alpha - \frac{3}{2}, 2\alpha - \frac{1}{2}\right]$. With non-selective disclosure, we have

$$U = 2\alpha - d_1 - E[d_2] = 2\alpha - d_1 - \frac{1}{2},$$

with ex-ante distribution

$$\begin{aligned} N(U) &= \Pr\left(2\alpha - d_1 - \frac{1}{2} \leq U\right) \\ &= U + \frac{3}{2} - 2\alpha \end{aligned}$$

for $U \in \left[2\alpha - \frac{3}{2}, 2\alpha - \frac{1}{2}\right]$. Thus, comparing the respective expressions for the distributions of U, we can confirm that $N(U)$ and $S(U)$ are indeed single crossing, while $S_n(\widehat{U})$ dominates $N(U)$ in the sense of FOSD.