Persuasion through Selective Disclosure: Implications for Marketing, Campaigning, and Privacy Regulation*

Florian Hoffmann† Roman Inderst‡ Marco Ottaviani§

March 2014

Abstract

By collecting personalized data, firms and political campaigners (senders) are able to better tailor their communication to the preferences and orientations of individual consumers and voters (receivers). This paper characterizes equilibrium persuasion through selective disclosure of the information that senders acquire about the preferences of receivers. We derive positive and normative implications depending on: the extent of competition among senders, whether receivers make individual or collective decisions, whether firms are able to personalize prices, and whether receivers are wary of the senders’ incentives to become better informed. We find that privacy laws requiring senders to obtain consent to acquire information are beneficial when there is little or asymmetric competition among firms or candidates, when receivers are unwary, and when firms can price discriminate. Otherwise, policy intervention has unintended negative welfare consequences.

Keywords: Persuasion, selective disclosure, hypertargeting, limited attention, privacy regulation.

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

---

*This paper supersedes and generalizes the analysis contained in an earlier draft circulated under the title: “Hypertargeting, Limited Attention, and Privacy: Implications for Marketing and Campaigning”. Hoffmann gratefully acknowledges financial support from the ERC Advanced Grant 22992, Ottaviani from the ERC Advanced Grant 295835, and Inderst from the DFG through the Research Unit “Design & Behavior” (FOR 1371).

†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 and IGIER, 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.

1 Introduction

Firms and political candidates have traditionally had two distinct ways to persuade consumers and voters. First, they could 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 strategies with direct marketing and ground-game campaigning aimed at persuading single individuals or small groups. In order to implement this second strategy, firms and candidates could hire experienced salespeople or skilled campaigners to gather critical knowledge about their audiences, making it possible to tailor their messages using face-to-face contacts or canvassing.

Nowadays, the greater availability of personally identifiable data on the internet blurs the distinction between these two traditional communication strategies. 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, and credit history, as well as the personal likes and dislikes the individual shared on social networking sites.\(^1\) This practice, known as behavioral targeting or hypertargeting, combines features of remote broadcasting with features of personal selling/campaigning.\(^2\)

When conducting what might appear to be an impersonal transaction through the internet, a great deal of personal information may be used to finely target consumers and voters.

---

1Information 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
Concerns are often raised that some consumers and voters might suffer if they remain blithely unaware of the ability of firms and candidates to collect information and communicate selectively. 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. 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.

This paper proposes a model to evaluate the private and social incentives for information acquisition by a set of firms or candidates (senders) who then communicate selectively to a set of consumers or voters (receivers). In the absence of regulation, senders secretly choose whether to acquire better private information about receiver preferences and then attempt to persuade by selectively disclosing attributes of their horizontally differentiated offerings. Thus, our baseline model hinges on three key features:

1. Senders decide whether to acquire better private information about receiver preferences.

2. Receivers do not observe whether senders have become better informed.

3. Senders and receivers play a disclosure game resulting in the selective revelation of the information that senders privately acquired. For illustration purposes, we cast the presentation within two classic specifications of selective disclosure:

   (a) In the first specification, each sender can increase the probability of becoming informed about a receiver’s valuation of the sender’s offering; then, the sender

---


4See 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.
discloses the value to the receiver. As it is well known from Dye (1985), Jung and Kwon (1998), and Shavell (1994), a sender who learns that the receiver has low values pools with the sender who does not become informed.

(b) Alternatively, in the second specification—a continuous version of Fishman and Hagerty’s (1990) model, similar to Glazer and Rubinstein (2004)—each sender decides whether to observe the receiver’s valuation for two attributes of the sender’s offering. However, the scope of communication is naturally restricted by factors such as airtime and screen space, or simply by the receiver’s limited attention. Given this limited attention, senders are able to disclose one of the two attributes, and thus select strategically the attribute to disclose so as to increase the chance of a favorable decision.

Thus, departing from the baseline models of Grossman (1981), Milgrom (1981), and Milgrom and Roberts (1986), in both specifications non-disclosure of an attribute does not trigger complete unraveling, either because (a) the sender is possibly uninformed in the first specification or (b) the receiver has limited attention and, thus, the sender can only disclose a subset (one “attribute”) of the available information in the second specification.

Our model of persuasion through selective disclosure is geared to address the following positive and normative questions in the economics of information privacy: In which circumstances do senders profit from acquiring more information before selectively disclosing? Do receivers benefit as a result? What is the impact of policies aimed at protecting privacy, for example by requiring consumer consent? Our answers depend on the extent of competition among senders, on whether the outcome is determined by the collective

---

5. Senders might be unable to communicate all the attributes they know because of space or time constraints or simply because (too much) information “consumes the attention of its recipients” (Simon 1971). 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 (Desselin, 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.

6. 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.

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 using 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. Our model is flexible enough to allow for a full-fledged analysis of the case with multiple senders with independently distributed values. In contrast, Gentzkow and Kamenica (2012) characterize optimal
action of multiple receivers (in the application to voting), on whether firms are able to personalize prices (in the marketing application), and on whether receivers are wary of the senders’ incentives to become better informed.

The three key features of our model of equilibrium persuasion without commitment depart from optimal persuasion with commitment à la Rayo and Segal (2010) and Kamenica and Gentzkow (2011) where, instead, the sender (1') chooses the information structure in an entirely unconstrained fashion, (2') is able to fully commit to this choice, and (3') thus completely discloses the information acquired. While the sender always (weakly) profits from optimal persuasion with full commitment, our analysis shows that equilibrium persuasion without commitment hurts the sender when the receiver’s outside option is relatively unattractive. As in the case of optimal persuasion, wary receivers benefit from equilibrium persuasion in the baseline model with individual decisions at given prices; however, receivers can be hurt when receivers make a collective decision as in our application to voting, when prices are personalized as in our marketing application, or when receivers are unwary.

In our model, senders are able to control receivers’ information indirectly by becoming better informed, our feature (1), and then by disclosing this information selectively, our feature (3), rather than directly and truthfully as in the literature on information control in markets à la Lewis and Sappington (1994), Johnson and Myatt (2006), and Gauza and Penalva (2010). In addition, senders in our model cannot commit to the information persuasion with multiple senders, when the three features of our equilibrium persuasion framework are not satisfied. Bhattacharya and Mukerjee (2013) analyze strategic disclosure by multiple senders who share the same information; in our horizontal-differentiation model, instead, senders are endogenously informed about the values of their offerings, which are independently distributed.

The setting with voting relates to the problem of persuading a group to take a collective decision considered by Caillaud and Tirole (2007); in our model voters cast their ballot simultaneously rather than sequentially. A political economy literature has focused on the welfare economics of contribution limits for financing campaign advertising; see, for example, Coate (2004) and the survey by Prat (2007); our model abstracts away from direct costs of campaign advertising.

Our analysis of the model with personalized pricing is also related to Gauza and Penalva’s (2010) work on the incentives for information provision in a second-price auction.

The analysis of the model with unwary receivers is related to work in the behavioral economics literature. 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. Unwary consumers in our second specification, 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 this model. 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).

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

See also Kamenica, Mullainathan, and Thaler (2011) for a discussion of situations in which firms
structure, our feature (2), and so can fall victim to their own incentives to secretly acquire more information; in these cases regulation can help senders achieve commitment.

To preview, we find 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 when these three conditions are not met, that is, when few competitors exploit unwary consumers through personalized pricing. 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.

Our analysis challenges the traditional distinction between persuasive and informative advertising.\textsuperscript{14} In our model, advertising is at once informative and persuasive, both on the way to the equilibrium—given that improved information shifts upward consumer demand, holding fixed consumer beliefs about information quality—and in equilibrium—given that even wary receivers end up being persuaded by the selective information they are given. An important twist is that the persuasion that results in equilibrium may actually hurt senders when competition is limited; in these circumstances senders benefit from committing not to acquire and disclose information.

We set the stage by introducing the impact of information on selective disclosure (Section 2). We proceed by analyzing our baseline model of equilibrium persuasion with a single receiver who is rational and faces fixed prices (Section 3). We then put equilibrium persuasion to work by applying it to campaigning (Section 4) and marketing with personalized pricing (Section 5). After revisiting all of our results for the case with receivers who are unwary of the senders’ incentives (Section 6), we conclude by summarizing the main insights and discussing open questions (Section 7). We collect the proofs in Appendix A and present additional material in Appendix B and Appendix C.

2 Selective Disclosure

At the heart of our analysis is a game of information acquisition and communication between senders and receivers. Senders attempt to persuade receivers to accept their offerings by tailoring their communication on the basis on what they learn about the preferences

\textsuperscript{14}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.
of any given receiver. Senders first acquire information about receiver preferences, where \( s_m \in \{L, H\} \) denotes that sender \( m \)'s information strategies involves less or more information. Communication then takes the form of disclosure, whereby senders truthfully reveal part of the information they have previously acquired. Thus, in line with the disclosure literature, the information that is transmitted is verifiable, for example because mendacious statements result in prohibitive losses of reputation or liability.

Disclosure affects a receiver’s perceived valuation of a sender’s offering, relative to the alternative option given by the best among the offerings of all the other senders. A given receiver will accept sender \( m \)'s offering if and only if the corresponding perceived utility \( U_m \) is sufficiently high, relative to the alternative option. We focus throughout on settings with horizontal differentiation where the values of the offerings are independent across senders and receivers. Given that a sender’s disclosure strategy does not affect the value of the offerings by the other senders, the analysis relies on the impact of the information strategy of a single sender on the perceived utility of each individual receiver. The main object of our analysis is the \textit{ex-ante} distribution of a receiver’s perceived utility from the offering of sender \( m, G_m(U_m) \), and how it is affected by the sender’s information strategy.

**Definition 1 (More Selective Disclosure).** We say that sender \( m \)'s information strategy \( s_m = H \) leads to “more selective disclosure” than \( s_m = L \) when the following holds for the \textit{ex-ante} distribution of a receiver’s perceived utility \( G_m(U_m) \):

(i) When the receiver believes \( s_m = H \), the resulting shift from \( G_m(U_m) = N_m(U_m) \) (for \( s_m = L \)) to \( G_m(U) = I_m(U_m) \) (for \( s_m = H \)) represents a Single-Crossing, Mean-Preserving Spread: There exists \( \tilde{U}_m \) so that for all \( U_m \in \text{interior of at least one support} \) it holds that \( I_m(U_m) > N_m(U_m) \) when \( U_m < \tilde{U}_m \) and \( I_m(U_m) < N_m(U_m) \) when \( U_m > \tilde{U}_m \).

(ii) When the receiver believes that \( s_m = L \), the resulting shift from \( G_m(U_m) = N_m(U_m) \) (for \( s_m = L \)) to \( G_m(U) = \tilde{I}_m(U_m) \) (for \( s_m = H \)) is such that \( \tilde{I}_m(U_m) \) dominates \( N_m(U_m) \) in the sense of strict \textit{First-Order Stochastic Dominance}.

That the shift in case (i) must be mean preserving follows immediately from Bayesian updating. The single-crossing of \( I_m(U_m) \) and \( N_m(U_m) \) implies that increased selectivity of disclosure moves probability mass away from the center and into the tails of the distribution. Hence, as disclosure becomes more selective it induces greater variability in receivers’ perceived utility. According to property (ii), when the receiver is unaware of the sender’s more selective information strategy, a better informed sender is able to induce in expectation a more favorable perception through selective disclosure. In our baseline analysis, a receiver will remain unaware, as in case (ii) of Definition 1, only off equilibrium; Section 6 extends the analysis to unwary receivers who remain unaware throughout.
In the remainder of this section we motivate Definition 1 in the context of two classic specifications of selective disclosure models. In both specifications, acquisition of more information about the preferences of an individual receiver results in more selective disclosure according to Definition 1. First, we consider the most immediate case of selective disclosure in which a sender decides whether or not to disclose information at all, depending on the what the sender has learned about the preferences of the receiver. We refer to this first specification as selective non-disclosure. But even when the sender always discloses some information, disclosure may still be selective when the sender can decide which information to disclose, provided that the sender has learned more than one piece of information about a receiver’s preferences. We refer to this second specification as selective targeted disclosure. We briefly introduce these two specifications in turn, and relegate additional analytical details to Appendices B and C. At the end of this section we also provide some additional general comments on Definition 1.

**Selective Non-Disclosure.** Denote the receiver’s true utility from alternative $m$ by $u_m$, which is independently distributed across all $M$ alternatives with logconcave prior distribution $F_m(u_m)$. For ease of exposition only, we stipulate a bounded support $[\underline{u}_m, \overline{u}_m]$ in the main text. Denote the expectation of $u_m$ by $E[u_m]$. Sender $m$ learns about $u_m$ with probability $\theta_m$ and then decides whether or not to disclose the respective information to the receiver.$^{15}$ The equilibrium disclosure strategy of sender $m$ is then characterized by a threshold value: When the receiver knows $\theta_m$, the sender only discloses when $u_m \geq u_d(\theta_m)$, which solves

$$u_d(\theta_m) = (1 - \theta_m)E[u_m] + \theta_mE[u_m | u_m \leq u_d(\theta_m)].$$

Uniqueness follows given that logconcavity of $F_m(u_m)$ implies that $dE[u_m | u_m \leq u']/du' \leq 1$. At $u_d(\theta_m)$ the disclosed true utility is equal to the receiver’s expected utility when there is no disclosure, as given by the right-hand side of (1).

Take two values $0 < \theta^L_m < \theta^H_m < 1$, so that $N_m(U_m)$ (for $s_m = L$) is generated with $\theta_m = \theta^L_m$, while $I(U_m)$ for $s_m = H$ is generated with $\theta_m = \theta^H_m$.$^{16}$ The single-crossing property of Definition 1(i) with crossing point $\tilde{U}_m = u_d(\theta^L_m)$, established in Appendix B, arises from two forces. First, the shift of mass to the upper tail follows as with an increase in $\theta_m$ it is also more likely that the sender observes and then discloses higher values of $u_m$. Second, the shift of mass to the lower tail is due to the downward shift in the lower

$^{15}$When the sender remains uninformed, we may stipulate that he either cannot disclose the respective information or that it is not optimal to do so. Otherwise, there is simply no scope for selective non-disclosure.

$^{16}$It is worthwhile to note that when the sender would always observe some signal about a receiver’s preferences, albeit possibly only a noisy one, full unraveling would necessarily result.
Figure 1: **Comparison of Distributions for Selective Non-Disclosure.** The left-hand panel depicts the shift from $N_m(U_m)$ to $I_m(U_m)$ (bold curve) with a uniform distribution, $F_m(u_m) = u_m$ and $\theta^L = \frac{1}{4}$, $\theta^H = \frac{3}{4}$. The right-hand panel depicts the shift from $N_m(U_m)$ to $\tilde{I}_m(U_m)$ (dotted curve).

support, $u_d(\theta^H_m) < u_d(\theta^L_m)$: As non-disclosure is now discounted more strongly by the wary receiver, it is optimal for the seller to disclose also lower values $u_m$. The left-hand panel in Figure 1 illustrates the transition from $N_m(U_m)$ to $I_m(U_m)$ for an example where $F_m(u_m)$ is uniformly distributed over $[0, 1]$.

Finally, to derive the distribution $\tilde{I}_m(U_m)$, suppose that the sender chooses $\theta_m = \theta^H_m$ but that the receiver is unaware of this (and thus still believes that $\theta_m = \theta^L_m$). Optimally, the sender still applies the cutoff $u_d(\theta^L_m)$, but, when now there is no disclosure, the receiver’s perceived value $U_m$ differs from the true conditional expected value by an amount equal $(\theta^H_m - \theta^L_m)[E[u_m] - E[u_m \mid u_m \leq u_d(\theta^L_m)]].$ The right-hand panel in Figure 1 illustrates the FOSD shift when $G_m(U_m)$ changes from $N_m(U_m)$ to $\tilde{I}_m(U_m)$, again for the uniform example. Appendix B shows that under selective non-disclosure the single-crossing property of $N_m(U_m)$ and $I_m(U_m)$ (Definition 1(i)) and the FOSD shift from $N_m(U_m)$ to $\tilde{I}_m(U_m)$ (Definition 1(ii)) hold generally for any logconcave distribution $F_m$.

**Selective Targeted Disclosure.** In the case of selective non-disclosure that we just discussed, full unraveling is avoided because a receiver always expects the sender to remain uninformed with positive probability. However, disclosure may still be selective and depend on the receiver’s individual preferences also in the case where a sender always discloses
some information, provided that the sender learns the receiver’s preferences. To model such selective targeted disclosure in the simplest possible way, we suppose that each sender has information about two different attributes of the offering, \( i = 1, 2 \). The receiver, if she knew both attributes, would learn two values \( u^i_m \), resulting in the true valuation

\[
 u_m = \sum_{i=1,2} u^i_m, \tag{2}
\]

in the spirit of Lancaster (1966). Valuations \( u^i_m \) are independently distributed across senders \( m \) and attributes \( i \) according to the distribution function \( F_m(u^i_m) \), which is again logconcave and has an atomless density \( f_m(u^i_m) \) over \([u^1_m, u^2_m]\). (Our slight abuse of notation across the two applications is intentional to stress similarities.) Though we will be more explicit about this below, in our application to marketing campaigns, \( i = 1, 2 \) may literally represent attributes of a product. The characteristics of a given attribute may then represent a good match for one receiver but not for another receiver.\(^{17}\) In Appendix C we also consider the case where the receiver places different weights on the utilities \( u^i_m \) in (2).

What prevents full unraveling with targeted selective disclosure is the limitation that the sender can only disclose one attribute, but not both. Our key motivation for this is a receiver’s limited attention, albeit in practice such a restriction also arises naturally from limitations of time and (advertising) space (see the discussion in the Introduction). Given that the same distribution \( F_m(u^i_m) \) applies to both \( u^1_m \) and \( u^2_m \), we can stipulate that when he is not informed about preferences, sender \( m \) always discloses the same attribute \( d_m \in \{1,2\} \) (disclosure strategy \( s_m = L \)). In this case, the receiver’s perceived utility will be \( U_m = u_{d_m} + E[u^i_m] \), so that the ex-ante distribution \( G_m(U_m) \) equals \( N_m(U_m) = F_m(U_m - E[u^i_m]) \). Instead, an informed sender (\( s_m = H \)) will disclose the attribute with the highest “fit”, \( d_m = \arg \max_{i=1,2} u^i_m \).\(^{18}\) The receiver, when aware of this, will discount the value of the non-disclosed attribute and obtain perceived utility

\[
 U_m = u_{d_m} + E[u^i_m \mid u^i_m \leq u_{d_m}].
\]

This gives rise to a unique and increasing value \( u_d(U_m) \), which retrieves the value that the receiver must have learned for a given perceived utility \( U_m \) to arise. The perceived utility

\(^{17}\)Then, an alternative interpretation is that disclosure by a firm allows a consumer to learn the distances between the product’s true characteristics and her own preferred characteristics. In an earlier version of this paper, using a Salop circle for each attribute, we have explicitly disentangled the actual location of product attributes from the preferred location according to the preferences of a particular consumer. The properties of Definition 1 then still hold despite the fact that the distribution of utility does not necessarily inherit the properties of the resulting distribution of distances.

\(^{18}\)A strategy where the sender does not disclose any attribute, but where the receiver knows that he is informed, would also not arise in equilibrium due to a standard unraveling argument.
Figure 2: Comparison of Distributions for Selective Targeted Disclosure. The left-hand panel depicts the shift from $N_m(U_m)$ to $I_m(U_m)$ (bold curve) with a uniform distribution, $F_m(u^i_m) = u^i_m$. The right-hand panel depicts the shift from $N_m(U_m)$ to $\tilde{I}_m(U_m)$ (dotted curve).

of a receiver who is aware that the sender discloses selectively the attribute with the highest value is distributed according to $I_m(U_m) = F^2_m(u_d(U_m))$. Instead, the perception about the non-disclosed attribute remains unchanged if the receiver is unaware that the sender discloses selectively, so that $U_m = u_{d_m} + E[u^i_m]$ with distribution $\tilde{I}(U_m) = F^2_m(U_m - E[u^i_m])$.

The left-hand panel of Figure 2 depicts the resulting shift from $N_m(U_m)$ to $I_m(U_m)$, for the example with a standard uniform distribution $F_m(u^i_m) = u^i_m$, as in Figure 1. That targeted selective disclosure, when the sender knows the receiver’s preferences, increases probability mass at the lower tail is again intuitive from the respective downward shift in the support of $U_m$: $N_m(U_m)$ has lower support $u_m + E[u^i_m]$, while that of $I_m(U_m)$ is $2u_m$ (with the lowest perceived value $U_m = 2u_m$ realized when $u_{d_m} = u_m$ is disclosed). At the upper tail of the distribution two conflicting forces are now at work under selective targeted disclosure: on the one hand, it is more likely that higher values of $u^i_m$ are disclosed when $s_m = H$ rather than $L$; on the other hand, the perceived value of the non-disclosed attribute is now discounted by the receiver, according to $E[u^i_m \mid u^i_m \leq u_{d_m}]$. We can show that with logconcavity of $F_m(u^i_m)$ the first effect dominates at the upper tail, as this sufficiently bounds the change in the conditional expected value of the non-disclosed attribute. The left-hand panel of Figure 2 illustrates the single crossing property according
to Definition 1(i) when $F_m$ is uniform.\textsuperscript{19} The right-hand panel illustrates that when the receiver is unwary, the resulting distribution $\tilde{F}_m(U_m)$ still FOSD dominates $N_m(U_m)$, which follows intuitively from the receiver’s inflated expectation about the undisclosed attribute.

We conclude the introduction of selective disclosure with some additional general observations. In light of our application to selective non-disclosure, we allow the distributions $G_m(U_m)$ to have an arbitrary (countable) number of mass points, though it is convenient to suppose that the respective support is convex and bounded. Further, we note that our analysis extends to the case in which the receiver’s utility depends on additional determinants which may be privately known to the receiver.\textsuperscript{20} To see this, suppose that the receiver’s total perceived utility under option $m$ is equal to $U_m$ plus some additional, independently distributed component $V_m$, i.e., equal to $\tilde{U}_m = U_m + V_m$, now with respective distribution $\tilde{G}_m(\tilde{U}_m)$. In particular, as long as $V_m$ has a logconcave density, when $G_m$ undergoes a (single-crossing) mean-preserving spread, the same holds for $\tilde{G}_m$, so that Definition 1 now applies to $\tilde{U}_m$.\textsuperscript{21}

3 Equilibrium Selective Disclosure

Model. Let $M \geq 2$ be the set, as well as the number, of the respective senders and their preferred actions, a different one for each sender. We assume throughout that values of the offerings are independently distributed across senders and across receivers. At $t = 1$ senders choose their information strategy $s_m$. Absent regulation, which is considered below, we have $s_m \in \{L, H\}$. The baseline assumption is that the choice of $s_m$ is an unobservable hidden action, which is again natural in the absence of regulation. At $t = 2$, senders can disclose information to the respective receiver. At $t = 3$, the receiver takes action $m \in M$ that results in the highest perceived utility, $U_m$, and randomizes with equal probability in case of a tie. With respect to the payoffs that are then realized, for the baseline specification of the model we only need to specify that each sender $m$ is strictly better off when the receiver chooses their offering $m$.\textsuperscript{22} Section 4 explicitly embeds our baseline analysis into a game of voting, while Section 5 extends the model to allow for personalized pricing. The receiver in the baseline is fully rational; Section 6 investigates

\textsuperscript{19}Appendix C also considers distributions with unbounded support.
\textsuperscript{20}All results generalize apart from the results with personalized pricing derived in Section 5), because the addition of receiver heterogeneity could interfere with optimal pricing.
\textsuperscript{21}Cf. e.g. Theorem 4 in Jewitt (1987).
\textsuperscript{22}Given that the distributions of preferences across alternatives are independent, we also need not specify whether sender $m$ receives a different (or the same) payoff when alternatives $m' \neq m$ or $m'' \neq m$ are chosen.
the case with unwary receivers.

**Receiver’s Preferences** Before we proceed to analyze the equilibrium, it is helpful to consider first the preferences of a receiver. The receiver realizes \( U^{(1)} = \max_{m \in M} U_m \). Suppose the receiver is aware of the strategies \((s_m)\) of all senders. Then, pick some \( m \in M \) and denote by \( G^{(1:M\setminus m)}(U^{(1:M\setminus m)}) \) the distribution of the maximum over all remaining senders, so that the receiver’s expected utility can be written as the following Lebesgue integral

\[
E[U^{(1)}] = \int \left[ \int \max \left\{ U^{(1:M\setminus m)}, U_m \right\} dG^{(1:M\setminus m)}(U^{(1:M\setminus m)}) \right] dG_m(U_m). \tag{3}
\]

As the expression in rectangular brackets is a convex function of \( U_m \), it is higher after a mean-preserving spread in \( G_m(U_m) \). From Definition 1 we thus have:

**Lemma 1** A receiver benefits when \( G_m(U) \) switches from \( N_m(U_m) \) to \( I_m(U_m) \).

The intuition for Lemma 1 is straightforward. For each alternative \( m \) the receiver has a binary decision to make, namely whether to take this alternative or not. Once the receiver decides against this alternative, the actual value of the respective utility is inconsequential as the receiver realizes, instead, the value of her next best alternative. Each alternative \( m \) thus represents an option for the receiver, whose expected value increases with a mean-preserving spread under \( I_m(U_m) \).

However, the paper analyzes three extensions of the model in which receivers will no longer unambiguously benefit from increased information acquisition by senders about their preferences:

(i) When decision-making is collective (Section 4);

(ii) When selective disclosure is combined with personalized pricing (Section 5); and

(iii) When a receiver (irrationally) remains unwary of senders’ information (Section 6).

**Senders’ Preferences and Equilibrium.** The senders’ incentives to become better informed are more subtle, already in this baseline model. Denote for a given realization of \( U_m \) the likelihood with which the receiver will choose alternative \( m \) by \( w_m(U_m) \) (“winning”), so that ex-ante alternative \( m \) is chosen with probability

\[
q_m = \int w_m(U_m) dG_m(U_m).
\]

\(^{23}\) It can be written as \( U_m + \int_{U_m}^{\infty} [U^{(1:M\setminus m)} - U_m] dG^{(1:M\setminus m)}(U^{(1:M\setminus m)}) \), which is a.e. differentiable with first derivative \( G^{(1:M\setminus m)}(U_m) \), which is increasing.
The following is an immediate implication of the fact that $\hat{I}_m(U_m)$ dominates $N_m(U_m)$ in the sense of strict First-Order Stochastic Dominance (FOSD) and that $w_m(U_m)$ is non-decreasing.\(^{24}\)

**Proposition 1** Suppose a receiver is not aware that sender $m$ discloses more selectively, so that $G_m(U_m)$ switches from $N_m(U_m)$ to $\hat{I}_m(U_m)$. Then, the only equilibrium is that where all senders who can do so will choose to disclose more selectively ($s_m = H$).

Suppose that the receiver believes that $s_m = L$. Then, however, according to part (ii) of Definition 1, the sender has a strict incentive to disclose more selectively ($s_m = H$), because this induces a favorable upward shift of the distribution of the receiver’s perceived utility from $G_m(U_m)$ to $\hat{I}_m(U_m)$. When all senders can do so, in equilibrium the receiver should thus anticipate $s_m = H$ for all $m \in M$.

Depending on circumstances, receivers may, however, observe a sender’s attempt to become better informed about receiver preferences in order to disclose more selectively. This is so, in particular, when information acquisition requires either the direct cooperation or at least the consent by receivers. In this case, matters are more nuanced, as we discuss next.

To gain intuition about the incentives at play, consider the case with a single sender and—only for the purpose of this initial illustration—assume that the receiver can only choose between alternative $m$ and a single “outside option” of known (reservation) value $R$. The receiver would then choose the sender’s preferred alternative with probability $1 - I_m(R)$ when the sender discloses (more) selectively, rather than with probability $1 - N_m(R)$. Given the single-crossing property in part (i) of Definition 1, the probability that the receiver accepts the sender’s offering is, thus, strictly higher under (more) selective disclosure if and only if $R$ lies to the right of the intersection of $I_m(U_m)$ and $N_m(U_m)$, while otherwise it is strictly lower. Intuitively, selective disclosure by a better informed sender results in a transfer of probability mass of the receiver’s perceived utility into the tails. Thus, it becomes more likely that the receiver chooses alternative $m$ when this is a priori rather unlikely; this situation is akin to Johnson and Myatt’s (2006) “niche market”.

In our model, when considering alternative $m$, the outside option of the receiver is the maximum of the perceived utility from all other alternatives, $U^{(1:M\setminus m)}$, rather than a fixed reservation value as in the case discussed in the previous paragraph. Consequently, we

\(^{24}\)The latter can be shown from an argument by contradiction. When all distributions $G_{m'}(U_{m'})$ are atomless, it is straightforward to derive $w_m(U_m) = \prod_{m' \in M \setminus m} G_{m'}(U_{m})$, as then the likelihood of a draw is zero. As we allow both for asymmetry across senders and for mass points of all $G_m(U_m)$, while the expression is straightforward to derive, it is generally quite unwieldy.
consider a shift in the attractiveness of a receiver’s outside option that results from an increase in competition, i.e., an increase in $M$.\textsuperscript{25} As there is more competition, it becomes increasingly likely from the perspective of each sender $m$ that a receiver’s best alternative option, $U^{(1:M,m)}$, takes on a high realization. Then, only high realizations of $U_m$ will ensure that the receiver still takes option $m$. What is thus key is that from Definition 1, $I_m(U_m)$ assigns more mass in the upper tail of the distribution, compared to $N_m(U_m)$.

To formalize this intuition, we must impose some restrictions as we increase $M$. In particular, it must be ensured that the addition of more senders indeed exerts sufficient competitive pressure on a given sender $M$. For instance, this would clearly not be the case when for all newly added senders $m'$ the mass of $G_{m'}(U_{m'})$ was confined to values $U_{m'}$ that are below the lower bound of the support of $G_m(U_m)$ (for a given choice $s_m$). We thus require that across all senders $m$, the distributions $N_m(U_m)$ have the same support, and that also the distributions $I_m(U_m)$ have the same support.\textsuperscript{26} Further, we stipulate that there is an arbitrarily large but finite set from which all senders (sender “types”) are drawn as we increase $M$, i.e., $G_m(U_m) \in \Omega$ for some finite $\Omega$. We will retain these specifications also subsequently whenever we consider the case of large $M$.

**Proposition 2** Suppose a sender’s choice of strategy $s_m \in \{L, H\}$ is observable. If competition is sufficiently intense (high $M$), any sender $m$ benefits when $G_m(U_m)$ switches from $N_m(U_m)$ to $I_m(U_m)$, so that the only equilibrium is the one where all senders who can do so disclose more selectively ($s_m = H$).

We can obtain a further characterization of the equilibrium, beyond that in Proposition 2, when we impose homogeneity across the distributions of receiver preferences for each option $m$ so that, in particular, $N_m(U_m) = N(U_m)$ and $I_m(U_m) = I(U_m)$. To see this, suppose first that all distributions are atomless, as is the case in our application of selective targeted disclosure. When all other senders, $M - 1$, disclose more selectively, then for any

\textsuperscript{25}Alternatively, one could consider a shift in the distribution of the receiver’s utility from a competing sender’s alternative. For $M = 2$ we can show the following for the preferred strategy of, say, sender $m = 1$: When he prefers $I_1(U_1)$ to $N_1(U_1)$ for a given choice of $G_2(U_2)$, then he strictly does so for any other distribution that dominates in the sense of the Monotone Likelihood Ratio Property. For an illustration, consider selective targeted disclosure with $u_1^* \propto \text{uniformly over } [0,1]$, as in Figure 2. For sender $m = 2$ take $G_2(U_2) = N_2(U_2)$ and suppose, only for now, that $u_2^*_2$ has a different support as it is distributed uniformly over $[0, \pi_2]$, with $0 < \pi_2 < 3$. Then $m = 1$ prefers to become informed and target his disclosure, resulting in $G_1(U_1) = I_1(U_1)$, if $\pi_2 > \frac{3}{4}$ and he prefers not to become informed if $\pi_2 < \frac{3}{4}$.

\textsuperscript{26}In terms of our two foundations, for instance, this condition is always satisfied when all values $u_m$ (where $u_m = u_m$ with selective non-disclosure and $u_m = u_m^*$ with selective targeted disclosure) have the same support, even though the respective distributions $F_m(u_m)$ may differ. Note that this condition does not require that, for given $m$, the respective distributions have the same support and that, despite having the same support, the respective distributions are the same across senders.
given sender we have \( w_m(U_m) = I^{M-1}(U_m) \). This allows us to show that there is a threshold value for \( M \) so that the outcome in Proposition 2 is indeed an equilibrium when \( M \) is above the threshold, but no longer an equilibrium when \( M \) is below the threshold. We can also extend this result to our application of selective targeted disclosure where all \( G_m(U_m) \) are no longer continuous.

**Proposition 3** In the context of both applications of selective non-disclosure and selective targeted disclosure, suppose that the distributions \( N(U_m) \) and \( I(U_m) \) in Definition 1 are identical across senders. Suppose also that a receiver becomes aware when sender \( m \) discloses more selectively. Then there is a finite threshold \( M' \) such that for all \( M \geq M' \) there exists an equilibrium where all senders disclose more selectively, \( s_m = H \), while for \( M < M' \) this outcome is not an equilibrium.

**Discussion: Implications for Regulation.** As discussed in the Introduction, absent regulation, consumers may not be able to control the extent to which firms collect personal data, such as past purchases that reflect also on consumer preferences for the present product. From Proposition 1, all firms would indeed want to disclose more selectively. From Lemma 1, consumers would strictly benefit from increased selective disclosure and, consequently, would be harmed by regulation that prohibits firms from collecting and using personal information. 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 become better informed about individual consumers’ preferences before disclosing selectively, 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. From Propositions 2 and 3 we know that a firm is more likely to benefit from such a commitment, as provided by regulation, when competition is limited (small \( M \)).

**Observation 1 (Regulation I)** Regulation that restricts the collection of personal data, from which senders can learn about receivers’ preferences, harms receivers in our baseline analysis. Even regulation that requires senders to obtain receiver consent is harmful when it enables senders to commit not to collect such data.

Observation 1 provides a benchmark case for analyzing regulation, to which we will return later in the paper after enriching the model.
4 Political Campaigning and Collective Decision-Making

“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.

We now extend the model to political campaigning by supposing that $M = 2$ candidates compete for voters. 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 $V$ of voters; the case with $V = 1$ coincides with our previous analysis. Denote by $U_m(v)$ voter $v$’s expected utility when candidate (or motion) $m$ wins, which now (for a given strategy $s_m$) is independently and symmetrically distributed across candidates. In terms of motivation, a candidate could generally be more or less forthcoming with respect to his views, i.e., more or less willing to disclose them (rather than not disclosing). With respect to selective targeted disclosure, candidates’ platforms could comprise issues on which a candidate’s stance can more or less coincide with the preference and political orientation of a particular voter.

Each voter casts the ballot for the candidate whose future decisions promise to deliver the highest utility value, and the candidate $m$ who obtains the highest number of voters wins the election. When $V^1$ is the number of votes cast for candidate $m = 1$ out of a total of $V$ votes, the ex-ante probability with which candidate $m = 1$ is elected is

\footnote{In recent field experiments, Dewan, Humphreys, and Rubenson (2014) and Kendall, Nannicini, and Trebbi (2013) find causal evidence of the effect of campaign information on intention to vote and electoral outcomes, respectively. For a view from the advertising industry, see Abse’s (2013) account of hypertargeting strategies pursued in the 2012 U.S. presidential election.}
\[ Q = 1 - \Pr[V^1 \leq (V - 1)/2] \]. Denote the likelihood of a vote for candidate \( m = 1 \) by \( q_1 = q \). (Of course, this will differ depending on candidates’ choice of strategy.) Then, given that \( V \) is odd, the expression for \( Q \) conveniently becomes

\[ Q = 1 - \sum_{k=0}^{\frac{V-1}{2}} \binom{V}{k} q^k (1 - q)^{V-k}. \]

Further, let \( y \) be the probability with which any given voter will become pivotal. Using again that \( V \) is odd, we have

\[ y = \left( \frac{V - 1}{V - 1/2} \right) q^{V-1/2} (1 - q)^{V-1/2}. \] (4)

Under collective decision-making, each voter’s ex-ante utility now comprises two terms. If the voter ends up not being pivotal, the vote does not influence the decision. As we now focus on the new effect arising from collective decision-making, we impose symmetry across candidates, implying in particular that when not being pivotal, a voter’s unconditional expected utility \( E[U] \) is independent on which candidate is chosen. However, 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 becomes

\[ E[U] + y \left( E[U^{(1)}] - E[U] \right). \] (5)

From Lemma 1, \( E[U^{(1)}] \), and thus the term in brackets, is strictly higher when one of the candidates discloses more selectively, and even more so when both candidates disclose more selectively. It is in this sense that our previous analysis can be applied. However, in this application to elections, a decision is no longer determined by the preferences of an individual receiver alone, as in our baseline case, but instead by the aggregate preferences of all receivers (voters). Therefore, we have to take into account a second effect. The strategies of candidates now also affect a voter’s likelihood \( y \) of becoming pivotal and, thereby, affect voter utility according to expression (5). This is what we explore in the rest of this section.

The likelihood of becoming pivotal decreases when, instead of \( q = 1/2 \), asymmetry in candidates’ strategies leads to \( q \neq 1/2 \), notably when one candidate alone, say \( m = 1 \), has a sufficiently sophisticated campaign team to collect the necessary information to “hyper-target” voters \( (s_1 = H) \). He will then indeed do so if this increases the likelihood of a vote in his favor, to \( q > 1/2 \). 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. We show next how this trade-off is always resolved unambiguously when the number of voters becomes sufficiently high.

Denote by $q_I$ and $y_I$ a voter’s probability of voting for $m = 1$ and for being pivotal when only one candidate, here candidate 1, can disclose more selectively ($s_1 = H$, $s_2 = L$). Denote by $q_N = 1/2$ and $y_N$ the corresponding probabilities when no candidate is better placed in this way ($s_1 = s_2 = L$). With a slight abuse of notation, $E_I[U^{(1)}]$ 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 (5), is higher when candidate 1 discloses more selectively if

$$\frac{y_I}{y_N} = \left[4q_I(1 - q_I)\right]^{\frac{V}{V+1}} > \frac{E_N[U^{(1)}] - E[U]}{E_I[U^{(1)}] - E[U]}.$$

(6)

While the right-hand side of the inequality (6) 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 $q_I \neq 1/2$. Hence, either condition (6) does not hold for all $V$, including $V = 3$, or there exists a cutoff value $\tilde{V}$ such that it is only for $V < \tilde{V}$ that a voter is better off when (only) candidate 1 discloses more selectively. 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 disclosure strategies (even though both $y_I$ and $y_N$ approach zero as $V \to \infty$).28

**Proposition 4** Suppose $V$ voters decide by majority rule over $M = 2$ ex-ante symmetric candidates. While voters are again strictly better off when both candidates disclose information more selectively ($G_m(U_m) = I(U_m)$), in contrast to the baseline analysis of individual decision making this is no longer unambiguously the case when only one candidate is in a position to do so. Then, while such asymmetric selectivity of disclosure still benefits each voter conditional on being pivotal, it decreases the likelihood with which individual voters become pivotal in the first place and, thereby, can decide the outcome based on their own preferences. There exists a cutoff $\tilde{V}$ on the number of voters such that voters are strictly worse off with asymmetric selectivity of disclosure when $V > \tilde{V}$, and better off when $V \leq \tilde{V}$.

**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

---

28With a uniform distribution over preferences, the threshold can be obtained explicitly. For instance, irrespective of the support, for the case with selective targeted disclosure we have $\tilde{V} = 153$. 
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.\footnote{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”.} 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 $z_v = z$.\footnote{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 (2011) for a model in which information sharing among voters reduces the incentives for information disclosure by candidates.}

5 Personalized Pricing

Scope for Personalized Pricing. In this section we consider in more detail the application of our model to marketing. Then a distinctive feature of our baseline analysis, where now senders represent firms and each receiver represents an individual consumer deciding which product to purchase, is that firms do not adjust prices individually, based on their knowledge of consumer preferences. Given that each firm offers all consumers the same product, even when it selectively gives them different information, such personalized pricing may be 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.\footnote{Also, price discrimination may be limited when consumers are concerned about fairness. Price (or rate) parity has become a major objective for firms, e.g., hotels, given the increasing transparency via online channels. 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.} These arguments motivate why there are circumstances under which our baseline analysis seems suitable. In other markets, however, because of transaction costs arbitrage may be of a lesser concern. 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.\footnote{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,
Firm Preferences with Personalized Pricing. 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 still purchases the product with the highest perceived utility $U^{(1)}$, and that, second, the price that the consumer pays is equal to the incremental utility relative to the second-highest such value. Consequently, a consumer realizes the second order-statistic, denoted by $U^{(2)}$. We first establish that with personalized pricing all firms prefer to disclose more selectively, now regardless of whether this is observed by consumers or not and irrespective of the intensity of competition (and thus in contrast to our previous findings without personalized pricing; see Propositions 2 and 3).

Recall our notation $U^{(1:M\setminus m)}$ for the highest expected utility over all other $M\setminus m$ firms. Then, the expected profit of firm $m$ is given by

$$\int \left[ \int \max \{U_m - U^{(1:M\setminus m)}, 0\} \ dG^{(1:M\setminus m)}(U^{(1:M\setminus m)}) \right] dG_m(U_m).$$

As the term in rectangular brackets is a convex function of $U_m$, to it is higher after a mean-preserving spread in $G_m(U_m)$. 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 the firm’s own product and the utility for the product of its closest rival, because the firm extracts exactly this difference. From an ex-ante perspective, the firm thus prefers a greater dispersion of $U_m$. As, trivially, the firm also benefits from a FOSD shift in $G_m(U_m)$, we have the following result:

**Proposition 5** With personalized pricing, provided it can do so, any given firm $m$ prefers to disclose more selectively, irrespective of whether this is observed by the consumer or not. Thus, with personalized pricing it is the unique equilibrium outcome that all firms that can do so disclose more selectively ($s_m = H$).

Consumer Preferences with Personalized Pricing. From consumers’ perspective, in stark contrast to the baseline case without personalized pricing, the effect of more selective disclosure depends now crucially on the degree of competition. As a starting point, take the case of a duopoly with $M = 2$. Here, the differences between the two cases are particularly stark. While without personalized pricing a consumer realized the maximum of the two expected utilities $U^{(1)} = \max\{U_1, U_2\}$, with personalized pricing

---

Villas-Boas (1999) and Acquisti and Varian (2005). As we abstract from this dynamic feature, our analysis will be quite different.

33Its derivative with respect to $U_m$ is $G^{(1:M\setminus m)}(U_m)$ and thus increases in $U_m$. 

20
the consumer now realizes the second-highest value, which for \( M = 2 \) is the minimum
\[ U^{(2)} = \min\{U_1, U_2\}. \]
The consumer is now strictly worse off when any of the presently considered two firms discloses more selectively. Formally, this can be seen in complete analogy to the argument to why without personalized pricing the consumer was strictly better off. More concretely, with personalized prices the consumer’s expected utility is
\[
E[U^{(2)}] = \int \left[ \int \min\{U_m, U_{m'}\} \, dG_{m'}(U_{m'}) \right] \, dG_m(U_m).
\]

As the expression in rectangular brackets is now a strictly concave function of \( U_m \), while it was a strictly convex function when without personalized pricing we applied the maximum, it is lower after a mean-preserving spread in \( G_m(U_m) \). The intuition for this is that as firm \( m \) discloses more selectively, a consumer’s updating makes firms, from an ex-ante perspective, more differentiated. This ensures that in expectation the (winning) firm with the highest perceived value can extract a higher price. As we show next, this negative effect of increased differentiation for consumers is, however, subdued when \( M \) increases, so that it becomes increasingly likely that each firm has a close competitor, in which case the efficiency benefits, as in Lemma 1, again accrue to the consumer.

More precisely, Proposition 6 establishes that regardless of what all other firms do, when there are sufficiently many firms a consumer strictly benefits when a particular firm discloses more selectively, so that \( I_m(U_m) \) instead of \( N_m(U_m) \) applies. In case all distributions are atomless, as in the case of selective targeted disclosure, then we have a particularly clear-cut result as to when a consumer benefits from a switch to \( I_m(U_m) \). As Proposition 6 shows, this is the case if and only if
\[
\int [N_m(U_m) - I_m(U_m)] \left[ G^{(2:M\setminus m)}(U_m) - G^{(1:M\setminus m)}(U_m) \right] \, dU_m > 0. \tag{7}
\]

Given our single-crossing condition for \( N_m(U_m) \) and \( I_m(U_m) \) (Definition 1(i)), in Proposition 6 we can sign expression (7) unambiguously to be positive whenever there are sufficiently many firms (high \( M \)). While we obtain expression (7) only after some transformations, it intuitively captures the fact that, from a consumer’s perspective, with personalized pricing the precise realization of \( U_m \) only matters when it falls between the first and second highest realizations of all other \( M - 1 \) utilities, which happens with probability \( G^{(2:M\setminus m)}(U_m) - G^{(1:M\setminus m)}(U_m) \). Expression (7) also makes it intuitive why in the following Proposition we need to exclude one particular case in which receivers do not benefit from a shift of probability mass into the tails of \( G_M(U_m) \).

\[ ^{34} \text{It can be written as } U_m - \int_{U_1}^{U_2} [U_m - U_{m'}] \, dG_{m'}(U_{m'}) \text{, which is a.e. differentiable with first derivative } 1 - G_{m'}(U_m) \text{, which is decreasing.} \]
**Proposition 6** When competition is sufficiently intense (high $M$), then even with personalized pricing a consumer strictly benefits when a given sender $m$ discloses more selectively ($s_m = H$), provided that at least one of the following conditions holds: Either the upper supports of $I_m(U_m)$ and $N_m(U_m)$ coincide or at least one other sender $m' \neq m$ chooses $s_{m'} = H$.

Proposition 6 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 6 we ask whether more information held by a single firm benefits the consumer.

### 6 Unwary Receivers

Definition 1 distinguishes between distributions $I_m(U_m)$ and $\hat{I}_m(U_m)$ for the case when sender $m$ discloses more selectively. So far the later case, $\hat{I}_m(U_m)$, only applied off equilibrium. In this section we consider the case of a receiver who remains unwary also in equilibrium. A receiver may generally not be aware of senders' capability to collect and use the respective personally identifiable data for selective disclosure. Further, a consumer may underestimate the skills of a salesperson as well as the underlying conflict of interest, and the same may happen for a voter. Given that the distribution $\hat{I}_m(U_m)$, as used in Definition 1(ii), concerns only a receiver's perceived utility, to derive implications for a receiver's welfare it is necessary to compute the receiver's true utility. To do so, we must specify the model of selective disclosure; nevertheless, all results we derive below hold for both selective non-disclosure and selective targeted disclosure.\(^{35}\)

**Baseline Analysis.** When a receiver remains unwary, the perceived value $U_m$ risks being inflated, given that from Definition 1(ii) $\hat{I}_m(U_m)$ is a FOSD shift of $N_m(U_m)$. With selective non-disclosure, inflation only results when sender $m$ chooses not to disclose information because otherwise the receiver observes the true utility. Both with selective non-disclosure and with selective targeted disclosure, the receiver's perceived value $U_m$ exceeds the true

\(^{35}\)In the selective disclosure specification, unwary receivers do not adjust for the the fact that the informed sender discloses the most favorable attribute. Thus, they are effectively "cursed", as in Eyster and Rabin (2005).
conditional expected value by $E[u_m'] - E[u_m | u_m \leq u_d]$ multiplied with the difference in the respective probabilities with which the sender becomes informed $(\theta_m^H - \theta_m^L)$. This inflation in the perceived value may distort decisions, making an unwary receiver worse off under $\tilde{I}_m(U_m)$ compared to $N_m(U_m)$. Less immediate, however, are two effects that work in the opposite direction, so that the senders’ better information—as a basis for more selective disclosure—end up being beneficial also for unwary receivers. One effect, which we will illustrate first, is that with incomplete information such inflated perceptions also reduce the error of inefficiently rejecting option $m$. The second effect works through competition when the perception of other senders’ options is equally inflated.

For an illustration of the first effect, take for selective targeted disclosure again the auxiliary case with a monopolistic sender and an outside option of known value $R$. Through biasing upwards an unwary receiver’s perceived utility relative to the known value $R$, the mistake of erroneously choosing the sender’s preferred option, even though $u^1 + u^2 < R$, evidently becomes larger when the sender discloses selectively. However, at the same time it becomes less likely that she erroneously decides against the sender’s preferred alternative, namely when actually $u^1 + u^2 > R$ holds. How these two errors trade off depends on the distribution $F(u^i)$. Interestingly, with a uniform distribution, as in Figure 2, it is straightforward to establish that these two errors exactly cancel out. When the receiver has to decide between some option $m$ and the next best alternative $m'$ whose perceived value is also inflated under competition, then these errors can disappear altogether, so that ultimately also an unwary receiver becomes unambiguously better off when (now all) senders disclose more selectively. Intuitively, under symmetry, when all senders choose the same disclosure strategy and when also the respective utilities are chosen from the same distributions, the respective distortions in perceptions fully cancel out. Then, for given choices of (non-)disclosure by all senders, an unwary receiver ends up making the same decision as a wary receiver. We state this result formally:

**Proposition 7** Suppose a receiver remains unwary when a sender discloses more selectively. While the impact for an unwary receiver of a particular sender switching to more selective disclosure is generally ambiguous, an unwary receiver is always strictly better off under symmetry, i.e., when preferences over all options $M$ are symmetric and when all

---

36 Note here that $u_m' = u_m$ applies with selective non-disclosure and $u_{m'} = u_{m'}^i$ with selective targeted disclosure.

37 With an exponential distribution, unwary consumers are in fact strictly better off when $G_m(U_m)$ switches from $N_m(U_m)$ to $\tilde{I}_m(U_m)$, even though in the case of monopoly their decision is biased. As these results for both the uniform and the exponential distribution hold for any choice of $R$, the observations extend to the case where $R$ is itself random and, thereby, also to when $R$ again represents the best alternative choice from $M - 1 > 0$ other senders (provided that the respective perception is not biased).
senders choose to disclose more selectively.

**Campaigning.** Recall that when there is collective decision-making, a voter’s expected utility is affected also by the probability with which she becomes pivotal. We show now how through this channel the presence of unwary receivers/voters also has a negative *externality* on other voters. For this recall first that we focused our analysis of campaigning on the case where voters have ex-ante symmetric preferences across candidates, but where an asymmetry is generated as only one candidate \(m = 1\) is sufficiently sophisticated to gather more information about voters and disclose (more) selectively. When a voter is unwary of this, from Definition 1(ii) it is immediate that this increases the likelihood with which she will vote for the respective candidate \(m = 1\), which we denote by \(\hat{q} > 1/2\). As this further reduces the likelihood with which any voter becomes pivotal, compared to the case where the considered voter is wary, we have:  

**Proposition 8** Consider the application to voting, where voters have a priori symmetric preferences over two candidates, and now suppose that \(V_u\) out of \(V\) voters remain unwary when one candidate discloses more selectively, while the other candidate cannot do so. Then, an increase in the number of unwary voters \(V_u\) decreases each voter’s probability of becoming pivotal, and thus also reduces each voter’s true expected welfare.

**Personalized Pricing.** Personalized pricing allows a firm to extract a consumer’s perceived incremental utility relative to her next best choice. If this perceived utility is inflated, this generates the potential for consumer exploitation. Again, this is most immediate in the auxiliary case where a monopolistic firm is not constrained by competition. A consumer who is unwary of a firm’s information will then realize with \(G(U) = \hat{I}(U)\) a true expected utility that is strictly below the reservation value. Thus, the consumer would be better off staying out of the market. However, in analogy to Proposition 6, competition again protects the consumer from such exploitation.

This effect comes out most clearly when firms are symmetric (cf. also Proposition 7). Then, recall first that if wary and unwary consumers receive the same information, they make the same choice. Interestingly, unwary consumers will now, however, pay a lower price than wary consumers if firms disclose more selectively. This follows because unwary consumers do not adjust expectations for the quality of information underlying

---

38Interestingly, in our application to selective targeted disclosure, we invariably obtain for this probability \(\hat{q} = \frac{2}{3}\), irrespective of the underlying distribution of preferences (cf. the proof of Proposition 8). Given ex-ante symmetry, when only one candidate can hypertarget in this way, an unwary voter’s probability to vote for this candidate would thus always increase by \(\frac{2}{3} - \frac{1}{2} = \frac{1}{6}\).
firms’ selective disclosure, which works towards reducing the perceived difference between the first-best and second-best alternative. To illustrate this formally, take the case of selective targeted disclosure. There, if $u^{(1)}$ is the highest disclosed fit (attribute) and $u^{(2)}$ is the second highest, an unwary consumer pays the price $\hat{p} = u^{(1)} - u^{(2)}$, given that the expectations about the non-disclosed attribute of either firm wrongly remain unchanged at $E[u^i]$. A wary consumer pays, instead, the strictly higher price

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

given that the term in brackets, equal to the difference in the updated conditional expectations for the attributes not disclosed by firms 1 and 2, is strictly positive. With selective non-disclosure, unwary consumers equally pay a strictly lower price than wary consumers whenever this is determined by the alternative of a firm that does not disclose. As in this case an increase in $\theta_m$ also increases the likelihood of disclosure, which (weakly) reduces the expected price and the likelihood of making the wrong choice, together with our previous results we have:

**Proposition 9** With personalized pricing and unwary consumers, selective disclosure has the potential for consumer exploitation, so that consumers realize a strictly lower expected true utility when participating in the market than when staying out of the market. However, this need not be the case. Instead, when competition is sufficiently intense (high $M$) and when firms’ options are symmetric, then with personalized pricing even an unwary consumer strictly benefits when all firms disclose more selectively.

We conclude by returning to our implications for regulation, again focused on the application to marketing and consumer–firm interaction. Compared to our baseline analysis, the introduction of personalized pricing and of consumer unwariness creates scope for inefficiencies and consumer exploitation—and, thereby, also creates scope for beneficial regulation from consumers’ perspective. However, we also showed how competition can ensure that efficiency benefits, as derived in our first observation in Lemma 1, dominate also from an unwary consumer’s perspective.

**Observation 2 (Regulation II)** If firms practice personalized pricing, and if consumers are unwary that senders are better informed about preferences when selectively disclosing information, consumer exploitation results, creating scope for beneficial regulatory intervention, such as restricting the collection of such personal data. However, under these same circumstances (with personalized pricing and unwary consumers), intervention can also harm consumers, especially so if competition is intense.
7 Conclusion

The greater availability of personally identifiable data opens up new opportunities for tailoring advertising messages to the perceived preferences of particular consumers or voters. Online marketing thus 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. Our simple model of equilibrium persuasion through selective disclosure can thus be applied to a broad range of settings:

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

- Second, consider a face-to-face interaction between a salesperson or campaigner and an individual consumer or voter. 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 and campaigning through communication channels that were previously anonymous but now allow for increased personalization, given the ability of senders to collect personally identifiable data on the internet. Based on an individual receiver’s profile, a firm may choose how to best use the limited amount of time or space to selectively convey the attributes of an offering.

We now summarize our main findings. In our baseline case with individual decisions at fixed prices, wary receivers benefit from an increase in information quality, given that they rationally adjust for the resulting increased selectivity of disclosure (Lemma 1). However, senders’ incentives to become better informed—as a basis of (more) selective disclosure—are more subtle. In the absence of policy intervention, we naturally assume that receivers do not observe the senders’ choice of information acquisition. Thus, for given expectations by the receivers, senders have an incentive to become better informed because, off-equilibrium, they would increase the chances that their offering is chosen (Proposition 1). Even though senders’ own incentives force them to become better informed, (more) selective disclosure ends up either benefitting or hurting them in equilibrium. Notably,
senders benefit from increased information when competition is intense, but they are hurt when competition is limited and their offerings are initially attractive (Propositions 2 and 3). In the latter case, policy intervention that makes information acquisition observable, for example by requiring consumer consent, hurts consumers because it allows senders to commit not to acquire information by not requiring consent.

Next, consider political campaigning by two candidates vying for a majority of voters. 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 utility conditional on this event. While the conditional utility always increases when more candidates are better informed about voters’ preferences before communicating selectively, at least when voters are wary, asymmetric abilities in information acquisition across candidates can tilt the vote shares of different candidates, thus reducing the probability that any given voter becomes pivotal. As a result, Proposition 4 finds that voters are made better off when any given candidate is better informed also in the asymmetric case where his competitor is not equally sophisticated, only as long as the number of voters is not too large. 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.

The introduction of personalized pricing changes the outcome of our baseline analysis 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 according to the perceived expected valuation of a particular consumer. In the presence of competition, perceived product differentiation matters. Selective disclosure based on better information dampens competition by increasing perceived differentiation, from an ex-ante perspective. Firms always benefit from such hypertargeting (Proposition 5), but the impact on consumer welfare depends on the intensity of competition (Proposition 6).

For policy purposes we check the robustness of our predictions to the baseline assumption that receivers are fully rational. When receivers remain unwary of the senders’ ability and incentives to become better informed, receivers may wrongly believe that senders are less informed and they do not properly discount the valuation for the adverse selection that is implicit in the fact that the disclosed attribute is the most favorable (in the case

---

39 Such price discrimination may only be feasible for services or low-value products, when customers or intermediaries have little scope for arbitrage.
of selective targeted disclosure) or that lack of disclosure is more likely to result with low realizations (in the case of selective non-disclosure). We find that:

- Unwary receivers can partly benefit from the increased information that sellers now *always* want to acquire, but may also lose because their choices become biased. However, Proposition 7 shows that in a fully symmetric setting, competition benefits unwary receivers particularly through the following mechanism. Given that unwary receivers have inflated perceptions about the offerings by all senders, the distortion in the processing of information obtained from one sender is compensated by a similar distortion of information obtained from competing firms, and exactly so under symmetry. Through this channel, competition eliminates bias and thus protects unwary consumers.

- For the model of campaigning, the presence of unwary voters has a negative externality on other voters (Proposition 8). When candidates differ in their capabilities to acquire information about individual voters’ preferences, an increase in the number of unwary voters decreases each voter’s probability of becoming pivotal, and thus also reduces each voter’s true expected welfare.

- In the case of personalized pricing, consumer ignorance about firms’ capabilities to acquire detailed information about customer preferences reduces differentiation and can thereby spur rivalry among firms for a particular consumer. As a consequence, with competition and price discrimination, consumer unwariness can even become a blessing. Overall, competition can protect unwary consumers from exploitation (Proposition 9).

Our results are obtained in a stylized model where senders can freely acquire and divulge information only about their own offerings in a private-value environment. First, we have abstracted away from learning about common values, which has been the focus of the literature on strategic voting (e.g., Feddersen and Pesendorfer 1996 and McMurray 2013). Second, an extension could allow senders to disclose information about their competitors. Third, information could be costly as in the law and economics literature on transparency,\(^40\) or it could be sold by an information broker as in Taylor (2004). Overall, our analysis reveals a twist to Shapiro and Varian’s (1997) policy of granting property rights over

\(^{40}\)When the prime purpose of information is to affect the distribution of surplus, the incentives to collect information may be too high (Hirshleifer 1971). 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.
information. We show that requiring consumer consent may allow firms to commit to abstain from selective communication, which would have benefited consumers.\footnote{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 (see 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


Appendix A: Proofs and Derivations

Proof of Proposition 2. For ease of exposition only we consider the choice of sender $m = 1$. Note, first, that when sender $m = 1$ switches from $N_1(U_1)$ to $I_1(U_1)$, the resulting difference in the likelihood $q_1$ is given by the following Lebesgue integral

$$
\int_{\hat{U}} w_1(U_1) d [I_1(U_1) - N_1(U_1)],
$$

(8)

where we choose the two values $\hat{U}$ and $\overline{U}$ large enough so that the (finite) supports of both $N(\cdot)$ and $I(\cdot)$ are contained in $[\underline{U}, \overline{U}]$. Recall from Definition 1(i) that $\tilde{I}_1$ denotes the point of intersection of $I_1(U_1)$ and $N_1(U_1)$ and take now some value $\tilde{U}_1 < \tilde{U} \leq \overline{U}$. Then, by extending (8) with $w_1(\tilde{U})$, it can be written as

$$
w_1(\tilde{U}) \left[ \int_{[\underline{U}, \tilde{U})} w_1(U_1) d [I_1(U_1) - N_1(U_1)] + \int_{[\tilde{U}, \overline{U}]} w_1(U_1) d [I_1(U_1) - N_1(U_1)] \right].
$$

(9)

We will next show that (9) is strictly positive for all sufficiently large $M$. In doing so, we will distinguish between the following three cases, which from Definition 1(i) exhaust all possibilities: (1) The two upper supports of $N$ and $I$ coincide, i.e., $\overline{U}_N = \underline{U}_I = \underline{U}$, and $I_1$ has no mass point at $\overline{U}_I$; (2) $\overline{U}_N = \underline{U}_I = \overline{U}$ but $I_1$ has a mass point at $\overline{U}_I$; (3) $\overline{U}_N < \overline{U}_I$.

Case 1: Define $G_m^\prime(U_1) = \lim_{U \to U_1} G_m(U)$ and note that monotonicity of $G_m(U)$ implies that

$$
\prod_{m' \in M \setminus 1} G_{m'}^\prime(U_1) \leq w_1(U_1) \leq \prod_{m' \in M \setminus 1} G_{m'}(U_1).
$$

Hence, for any two $U_1'' > U_1'$ it holds that

$$
\frac{w_1(U_1'')}{w_1(U_1')} \geq \prod_{m' \in M \setminus 1} \frac{G_{m'}^\prime(U_1'')}{G_{m'}(U_1')},
$$

(10)

From Definition 1(i) and the properties of Case 1, there exists $\hat{U} < \overline{U}$ so that the following holds. First, we have for $U_1 \geq \hat{U}$ that $dI(U_1) > dN(U_1)$. Second, defining for any $U_1 < \hat{U}$

$$
\varphi_\omega(U_1) = \frac{G_\omega(U_1)}{G_\omega(\hat{U})},
$$

we have $\varphi(U_1) = \max_{\omega \in \Omega} \varphi_\omega(U_1) < 1$. So, applying (10), it surely must hold for any $U_1 < \hat{U}$ that

$$
\frac{w_1(U_1)}{w_1(\hat{U})} \leq [\varphi(U_1)]^{M-1}.
$$

\footnote{Note that both $I_1(\cdot)$ and $N_1(\cdot)$ are right-continuous functions.}
That \( w_1(U_1)/w_1(\tilde{U}) \to 0 \) uniformly as \( M \to \infty \) follows then from the observations that \( \varphi(U_1) \in [0,1) \) and that \( U_1 \) is taken from a bounded set. As a consequence, if we now consider a sequence of markets, indexed by \( M \), we have

\[
\lim_{M \to \infty} \int_{[\underline{U}, \bar{U}]} \frac{w_{1,M}(U_1)}{w_{1,M}(\bar{U})} d[I_1(U_1) - N_1(U_1)] = 0. \tag{11}
\]

Making now use of \( \hat{U} < \underline{U} \), \( dI(\bar{U}_1) > dN(\bar{U}_1) \) for all \( \hat{U} \leq U_1 < \bar{U} \), as well as obviously

\[
\frac{w_1(U_1)}{w_1(\bar{U})} \geq 1
\]

over this range, there surely exists a value \( \varepsilon > 0 \) such that along any considered sequence we have

\[
\int_{[\hat{U}, \bar{U}]} \frac{w_{1,M}(U_1)}{w_{1,M}(\bar{U})} d[I_1(U_1) - N_1(U_1)] > \varepsilon. \tag{12}
\]

Substituting (11) and (12) into (9) proves the assertion for Case 1.

Case 2: With \( \bar{U}^N = \bar{U}^I = \bar{U} \) but \( I(U_1) \) having a mass point at the upper bound, we choose \( \hat{U} = \bar{U} \). From Definition 1(i), (12) is then again immediate. (In particular, in case also \( N(U_1) \) has a mass point at \( \bar{U} \), then this is strictly smaller.) Next, as \( \bar{U} \) is the upper bound of both supports, also the preceding argument for all \( U_1 < \hat{U} \) applies, so that also (11) holds.

Case 3: With \( \bar{U}^N < \bar{U}^I \), we set \( \hat{U} = \bar{U}^N \). The first part of the argument, as in Case 1, now applies as \( w_1(U_1)/w_1(\hat{U}) \to 0 \) holds, using convexity of the supports, for all \( U_1 < \hat{U} \). Hence, (11) again applies. We come next to the second part of the argument. Using monotonicity of \( w_1(U_1) \), we can write

\[
\int_{[\hat{U}, \bar{U}]} \frac{w_{1}(U_1)}{w_{1}(\bar{U})} d[I_1(U_1) - N_1(U_1)] \geq \left[ 1 - I_{\hat{U}}(\bar{U}^N) \right] - \left[ 1 - N_{\hat{U}}(\bar{U}^N) \right],
\]

which is strictly positive from Definition 1(i). Thus, also (12) holds along any considered sequence with \( M \to \infty \). This completes the proof. Q.E.D.

**Proof of Proposition 3.** Assume that \( N_m(U_m) = N(U_m) \) and \( I_m(U_m) = I(U_m) \) for all \( m \) and consider, first, the case without mass points (e.g. the case of selective targeted disclosure). For ease of exposition only we consider again the choice of sender \( m = 1 \). Then, when all \( M\setminus1 \) senders disclose more selectively, the difference in the likelihood \( q_1 \) that option \( m = 1 \) is chosen when also sender \( m = 1 \) does so is given by

\[
\Delta q_1(M) = \int w_1(U) d[I_1(U_1) - N_1(U_1)],
\]

with \( w_1(U) = I^{M-1}(U) \). Using \( Z(U) = N(U) - I(U) \), which by Definition 1(i) is, in the interior of the respective supports, strictly negative for \( U < \tilde{U} \) and strictly positive for
$U > \tilde{U}$, we can write after integration by parts, which we can apply as there are no mass points,

$$\Delta q_1(M+1) = \int_{\tilde{U}}^{U} Z(U) \frac{MI(U)}{(M-1)} dI^{M-1}(U) + \int_{\tilde{U}}^{\tilde{U}} Z(U) \frac{MI(U)}{(M-1)} dI^{M-1}(U)$$

$$> \frac{MI(\tilde{U})}{(M-1)} \Delta q_1(M).$$

This implies that whenever $\Delta q_1(M) \geq 0$ it must hold that $\Delta q_1(M+1) > 0$. Existence of a threshold $M' \geq 2$ follows then together with Proposition 2.

We turn now to the case where $G_m(U_m)$ can have a mass point. Note that we here restrict consideration to selective non-disclosure. Choose again $G_m(U) = I(U)$ for all $m \in M\setminus 1$ and recall that here $N(U)$ has an atom at $U = \tilde{U}$, while the single atom of $I(U)$ is at a strictly lower value, $\underline{U}$, which is also the lowest realization. Defining $j = I(\underline{U})$ we thus can write

$$\Delta q_1(M+1) = \int_{\underline{U}}^{\tilde{U}} Z(U) \frac{M}{(M-1)} I(U) dI^{M-1}(U) - \left( \frac{M}{M-1} \right) \frac{M-1}{M} j^M$$

$$> \frac{M}{M-1} I(\tilde{U}) \left( \int_{\underline{U}}^{\tilde{U}} Z(U) dI^{M-1}(U) + \int_{\tilde{U}}^{\tilde{U}} Z(U) dI^{M-1}(U) - \frac{M-1}{M} j^M \right)$$

$$= \frac{M}{M-1} I(\tilde{U}) \Delta q_1(M),$$

where we have used Definition 1(i) and, from $j = I(\underline{U})$, that $\frac{M}{M-1} \frac{M}{M+1} j < \frac{M}{M-1} \frac{M}{M+1} I(\tilde{U}) < \frac{M}{M-1} I(\tilde{U})$. The result then follows as before. Q.E.D.

**Proof of Proposition 6.** To simplify the exposition, without loss of generality we consider again the choice of firm $m = 1$ and thus a switch from $G_1(U_1) = N_1(U_1)$ to $G_1(U_1) = I_1(U_1)$. Note also that the case with $M = 2$ was already fully solved in the main text and that, presently, we are interested in the case for high $M$, which is why without loss of generality we can assume that $M \geq 3$. Denote by $v_1(U_1)$ a consumer's expected utility for given $U_1$, so that the respective difference in ex-ante utility is given by

$$\int_{\underline{U}}^{\tilde{U}} v_1(U_1) d [I_1(U_1) - N_1(U_1)]. \quad (13)$$

Next, note that $v_1(U_1)$ is continuous and almost everywhere differentiable with left-hand side derivative $\lim_{U_1 \to U_1^{-}} v'_1(U) = \lim_{U_1 \to U_1^{-}} G^{(2:M\setminus 1)}(U_1) - G^{(1:M\setminus 1)}(U_1) = \eta(U_1)$, where

$$\eta(U_1) = \sum_{m \in M\setminus 1} \left[ 1 - G_m(U_1) \right] \prod_{m' \notin \{1,m\}} G_{m'}(U_1).$$

36
Then, using
\[
\int_U \eta(U_1) [N_1(U_1) - I_1(U_1)] dU_1. \tag{14}
\]
This generalizes expression (7) in the main text, where for ease of exposition it was assumed that all distributions are continuous. In order to show (14) to be strictly positive for high \(M\), we distinguish two cases: 1) \(\tilde{U}_1 < \tilde{U}^N\) and 2) \(\tilde{U}_1 = \tilde{U}^N\).

Case 1: First, assume that \(\tilde{U}_1 < \tilde{U}^N\) and define \(\phi(U_1) = \eta(U_1)/\eta(\tilde{U}_1)\). Then, extending the expression in (14), for each \(U_1\), by multiplying and dividing with the term \(\eta(\tilde{U}_1)\), and noting that \(\eta(\tilde{U}_1) > 0\), a sufficient condition for (14) to be greater than zero is that
\[
\int_{[\tilde{U}_1, \tilde{U}]^c} [N_1(U_1) - I_1(U_1)] \phi(U_1) dU_1 + \int_{[\tilde{U}_1, \tilde{U}]} [N_1(U_1) - I_1(U_1)] \phi(U_1) dU_1 > 0. \tag{15}
\]
Next, define, using finiteness of \(\Omega\), for any \(U_1\), \(\zeta = \max_{\omega \in \Omega} [1 - G_\omega(U_1)]\) and \(\zeta' = \min_{\omega \in \Omega} [1 - G_\omega(\tilde{U}_1)] > 0\) and note that for \(U_1 < \tilde{U}_1\) we have
\[
\phi(U_1) \leq \left(\frac{\zeta}{\zeta'}\right) \frac{\sum_{m \in M \setminus 1} \prod_{m' \in \{1, m\}} G_{m'}^-(U_1)}{\sum_{m \in M \setminus 1} \prod_{m' \notin \{1, m\}} G_{m'}^-(\tilde{U}_1)}.
\]
Then, using \(\kappa(U_1) = \max_{\omega \in \Omega} \frac{G_\omega(U_1)}{G_\omega(\tilde{U}_1)} < 1\) for \(U_1 < \tilde{U}_1\), it holds that \(\phi(U_1) \leq \left(\frac{\zeta}{\zeta'}\right) \kappa^{M-2}(U_1)\).

From this, together with the observations that \(U_1\) is chosen from a bounded set and \(\kappa(U_1) \in [0, 1]\), it follows that \(\phi(U_1) \to 0\) uniformly as \(M \to \infty\). It then follows by the same arguments as in the proof of Proposition 2 that the first integral in (15) goes to zero as \(M \to \infty\). It remains to show that the second integral remains bounded away from zero (and positive). Clearly, for \(U_1 > \tilde{U}_1\) we have from Definition 1(i) that \(N_1(U_1) > I_1(U_1)\) such that it is sufficient to show that the weights \(\phi(U_1)\) do not go to zero on a set of positive measure. But this follows immediately from the fact that, for any \(\tilde{U}_1 < U_1 < \tilde{U}^N\) in the interior of all supports (Case 1), we have
\[
\phi(U_1) \geq \left(\frac{\zeta''(U_1)}{\zeta'(U_1)}\right) \frac{\sum_{m \in M \setminus 1} \prod_{m' \notin \{1, m\}} G_{m'}^-(U_1)}{\sum_{m \in M \setminus 1} \prod_{m' \notin \{1, m\}} G_{m'}^-(\tilde{U}_1)},
\]
where \(\zeta''(U_1) = \min_{\omega \in \Omega} [1 - G_\omega^-(U_1)] > 0\) and \(\zeta'(U_1) = \max_{\omega \in \Omega} [1 - G_\omega^-(\tilde{U}_1)]\). Then, using \(\kappa(U_1) = \min_{\omega \in \Omega} \frac{G_\omega(U_1)}{G_\omega(\tilde{U}_1)} > 1\), it holds that \(\phi(U_1) \geq \frac{\zeta''(U_1)}{\zeta'(U_1)} \kappa^{M-2}(U_1) > \frac{\zeta''}{\zeta'} > 0\).
Case 2: Here it holds that $\tilde{U}_1 = \overline{U}^N$. From Definition 1(i), this case can only arise when $\overline{U}^N < \overline{U}^I$ and $N_1$ has a jump at $\tilde{U}_1 = \overline{U}^N$. Then, assume, first, that $G_m = N_m$ for all $M \setminus m$, implying that $\eta(U_1) = 0$ for all $U_1 > \overline{U}^N$. Hence, we can write (14) as

$$\int_{[\overline{U}, \tilde{U}_1]} \eta(U_1)[N_1(U_1) - I_1(U_1)]dU_1 \leq 0.$$ 

A necessary condition for (14) to be strictly positive for high $M$ is that the set $M^* \subseteq M \setminus \{1\}$ where $G_{m^*} = I_{m^*}$ for all $m^* \in M^*$ is non-empty. We will show next that this is also sufficient. To see this take any $\tilde{U}_1 = \overline{U}^N < \tilde{U} < \overline{U}^I$, and extend the expression in (14), for each $U_1$, by multiplying and dividing with the term $\eta(\tilde{U}) > 0$. Then, in analogy to (15), a sufficient condition for (14) to be greater than zero is that

$$\int_{[\overline{U}, \tilde{U}_1]} [N_1(U_1) - I_1(U_1)] \phi(U_1)dU_1 + \int_{[\overline{U}, \tilde{U}_1]} [N_1(U_1) - I_1(U_1)] \phi(U_1)dU_1 > 0. \quad (16)$$

Clearly, the first integral in (16) still goes to zero as $M \to \infty$ by the same argument as in the proof of case 1. To show that the second integral is bounded away from zero (and positive), from $N_1(U_1) - I_1(U_1) > 0$ for all $\tilde{U} < U_1 < \overline{U}^I$, it suffices to show that the weights $\phi(U_1)$ do not go to zero on a set of positive measure. But, using $\zeta''(U_1) = \min_{m \in M^*}[1 - I_m(U_1)] > 0$ and $\zeta''(U_1) = \max_{m \in M^*}[1 - I_m(\tilde{U})]$, this follows immediately from

$$\phi(U_1) \geq \left(\frac{\zeta''(U_1)}{\zeta''(U_1)}\right) \frac{\sum_{m \in M^*} \prod_{m' \in M \setminus \{1, m\}} G_{m'}(U_1)}{\sum_{m \in M^*} \prod_{m' \in M \setminus \{1, m\}} G_{m'}(\tilde{U})} \geq \left(\frac{\zeta''}{\zeta''}\right) \nu^{M-2}(U_1),$$

where $\nu(U_1) = \min_{\omega \in \Omega} \frac{G_\omega(U_1)}{G_\omega(\tilde{U})} \geq 1$. Q.E.D.

**Proof of Proposition 7.** We will prove the result separately for the two cases with selective targeted disclosure and selective non-disclosure. Consider first selective targeted disclosure. In this case senders always disclose $u_{dm} = \max\{u_{m1}, u_{m2}\}$ irrespective of whether they face a wary or an unwary receiver. As both the true expected valuation, $u_{dm} + E[u_m' | u_m' \leq u_{dm}]$, as well as the one perceived by the unwary receiver, $u_{dm} + E[u_m']$, are strictly increasing in $u_{dm}$, and, by symmetry, for a given $u_{dm}$ constant over $m$, an unwary receiver’s decision rule is the same as that of a wary receiver, namely to choose the option where the respective disclosed value $u_{dm}$ is maximal. Thus, wary and unwary receivers realize the same expected utility when all senders disclose an attribute selectively. Hence, we conclude from Lemma 1 that unwary consumers, like wary consumers, are better off with selective disclosure.
Next, consider the case of selective non-disclosure. Under $G_m(U_m) = \hat{I}(U_m)$, with $s_m = H$ senders still disclose when $u_m \geq u_d(\theta^L)$. Also, given symmetry, when no firm chooses to disclose, the unwary receiver chooses one of the $M$ options at random, while when at least one sender chooses to disclose, she correctly chooses the alternative that gives her the highest true expected utility. The claim then follows immediately as the probability of disclosure increases for each sender if $\theta$ switches from $\theta^L$ to $\theta^H$. Q.E.D.

**Proof of Proposition 8.** Note first that the expected utility of a pivotal voter, no matter whether she is wary or unwary, is independent of the number of unwary voters $V_u$. Hence, $V_u$ can only affect expected voter welfare through its effect on the probability of becoming pivotal. Recall that we denote the probability that a wary voter elects candidate $m = 1$ by $q$ and the respective probability for an unwary voter by $bq$. As $\hat{I}(U)$ dominates $N(U)$ in the sense of strict FOSD according to Definition 1(ii), we must have that $\hat{q} > \frac{1}{2}$. We derive this next more explicitly for the selective non-disclosure and selective targeted disclosure and relate this to the case where the considered voter is wary, showing that $\hat{q} > q \geq 1/2$.

For selective targeted disclosure we have

$$\hat{q} = \int_{U}^{\pi E[u']}(1 - \hat{I}(U)) dN(U) = \int_{\pi E[u']}^{\pi E[u']} \hat{I}(U) N(U) dU$$

$$= \int_{\pi E[u']}^{\pi E[u']} 2f(U - E[u'])F^2(U - E[u']) dU = \left[\frac{2}{3} F^3(U - E[u'])\right]_{\pi E[u']}^{\pi E[u']} = \frac{2}{3},$$

while

$$q = \left[\frac{2}{3} F^3(u_d(U))\right]_{\pi E[u']}^{\pi E[u']} < \hat{q} = \frac{2}{3}. $$

Next, for selective non-disclosure we have

$$q = \frac{1}{2} \frac{\theta^H}{\theta^L} \left[1 - N^2(u_d(\theta))\right]$$

and

$$\hat{q} = q + \frac{1}{2} N(u_d(\theta^L))\hat{I}(u_d(\theta^L)).$$

We show finally that $\hat{q} > q \geq 1/2$ implies that the likelihood of becoming pivotal decreases with $V_u$. To see this, 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. 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 $V \setminus \{v, v'\}$ 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 \( q_v X + (1 - q_v) Y \). From \( \bar{q} > q \geq 1/2 \) it follows that \( X < Y \)\(^{43}\) and that \( q_v X + (1 - q_v) Y \) is larger for \( q_v = q \) than for \( q_v = \bar{q} > q \). Q.E.D.

\(^{43}\)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.
Appendix B: Selective Non-Disclosure

This appendix provides additional details for the case of selective non-disclosure. For the present results we can suppress the subscript \( m \) denoting the respective sender. Note now, first, that uniqueness of \( u_d(\theta) \), as given in (1), follows from

\[
\frac{dE[u | u \leq u']}{du'} \leq 1,
\]
which in turn is implied by logconcavity of \( F(u) \). This holds if and only if \( H(u') = \int_{u}^{\infty} F(u) du \) is logconcave, which is implied by logconcavity of \( F(u) \); see Bagnoli and Bergstrom (2005), in particular their Lemma 1 and Theorem 1. From (17) we also have that

\[
\frac{du_d(\theta)}{d\theta} = -\frac{E[u] - E[u | u \leq u_d(\theta)]}{1 - \theta \frac{dE[u | u \leq u_d(\theta)]}{du_d(\theta)}} < 0.
\]

To characterize \( N(U) \) and \( I(U) \), where \( \theta \) is known to the receiver, note that in either case, with \( \theta = \theta_L \) or \( \theta = \theta_H \), we have for \( u_d(\theta) \leq U \leq \overline{u} \) that \( G(U) = (1 - \theta) + \theta F(U) \), while at the lower bound there is a jump \( j(\theta) = G(u_d(\theta)) \), satisfying

\[
\frac{dj(\theta)}{d\theta} = \theta \frac{du_d(\theta)}{d\theta} f(u_d(\theta)) - (1 - F(u_d(\theta))) < 0,
\]
where we made use of (18). As we move from \( \theta = \theta_L \) to \( \theta = \theta_H \), single crossing, as in Definition 1(i), thus follows from (18), (19), and from the fact that \( G(U) \) is strictly decreasing in \( \theta \) (with slope \( F(U) - 1 \)) for all \( U > u_d(\theta) \). Note also that the single crossing point is at the lower support of \( N(U) \), \( u_d(\theta_L) \).

Finally, when the receiver is unwary of the switch to \( \theta = \theta_H \), we have for \( \tilde{I}(U) \) the following: It is distributed over \([u_d(\theta_L), \overline{u}]\) with \( \tilde{I}(U) = (1 - \theta_H^0 + \theta_H F(U) \) and a jump of size \( 1 - \theta_H^0 (1 - F(u_d(\theta_L))) \) at the lower bound. From this we have immediately that \( \tilde{I}(U) < N(U) \) holds strictly for all \( u \in [u_d(\theta_L), \overline{u}] \), i.e., at all points of their joint support apart from the upper bound. We have thus established also FOSD, as required by Definition 1(ii). Overall, we have the following result.44

Claim 1 (Appendix B). In the case of selective non-disclosure, \( N(U) \), \( I(U) \), and \( \tilde{I}(U) \) satisfy the respective properties of Definition 1.

Appendix C: Selective Targeted Disclosure

This appendix provides additional details for the case of selective targeted disclosure. Again, it is convenient to suppress the subscript \( m \). Take first the case of a uniform

44 All the results extend also to the case with an unbounded upper support \( \overline{u} \to \infty \).
distribution (as in Figure 2). With support \([\underline{u}, \overline{u}]\), \(N(U)\) and \(I(U)\) are given by

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

which clearly can have at most two intersections. It is easily verified that these occur at the shared upper bound of the support \(\overline{U} = \pi + E[u] = (3\pi + \underline{u})/2\) and at \(\tilde{U} = (3\pi + 5\underline{u})/4\), with \(\underline{u} + E[u] < \overline{U} < \pi + E[u]\). For completeness, note also that

\[
\hat{I}(U) = \left(\frac{2U - 3\underline{u} - \pi}{2(\overline{u} - \underline{u})}\right)^2.
\]

**Claim 1 (Appendix C).** When attribute values \(u\) are uniformly distributed, then the respective distributions of perceived utility (as given explicitly in (20) and (21)) satisfy Definition 1.

Next, we formally support the claim that logconcavity of \(F(u)\) is generally sufficient to ensure that \(I(U)\) has more mass than \(N(U)\) also in the upper tail, i.e., that there exists \(U' < \overline{U}\) such that \(N(U) > I(U)\) for all \(U \geq U'\). With a slight abuse of notation, denote for given realized value \(U\) the corresponding disclosed values respectively by \(u_N = U - E[u]\) and \(u_I = u_d(U)\), where at \(U = \overline{U}\) we have \(u_I = u_N = \overline{u}\). The densities of the two distributions are then, generally,

\[
n(U) = f(u_N), \quad i(U) = f(u_I) \frac{2F(u_I)}{1 + \frac{dE[u]u'|u' \leq u_I]} du_I}.
\]

Together with \(u_I = u_N\) at \(U = \overline{U}\), next to \(F(u_I = \overline{u}) = F(u_N = \overline{u}) = 1\), this yields \(n(\overline{U}) < i(\overline{U})\) if and only if, at this point, condition (17) holds (with \(u' = u_I\)). As noted in Appendix B, this is implied by logconcavity of \(F(u)\).

**Claim 2 (Appendix C).** \(I(U)\) has more mass also in the upper tail, i.e., \(I(U) < N(U)\) holds for all sufficiently high \(U < \overline{U}\).

As noted in the main text, we also establish Definition 1 for a case where the support of the distribution is unbounded, namely that where \(u\) is distributed exponentially with parameter \(\lambda\).

**Claim 3 (Appendix C).** When attribute values \(u\) are exponentially distributed (and thus have unbounded support), the then derived distributions \(N(U), I(U), \) and \(\hat{I}(U)\) still satisfy Definition 1.
**Proof.** As \( \bar{T}(U) \) always FOSD dominates \( N(U) \), we can restrict ourselves to a comparison of \( N(U) \) and \( I(U) \), which are given by \( N(U) = 1 - e^{1 - \lambda U} \) and \( I(U) = \left(1 - e^{-\lambda u_d(U)}\right)^2 \), where \( u_d(U) \) solves

\[
U = u_d + \frac{1}{\lambda} - \frac{e^{-\lambda u_d}}{1 - e^{-\lambda u_d}} u_d.
\]

(22)

For \( Z(U) = N(U) - I(U) \) we have

\[
Z(U) = -\exp\{ -\lambda u_d(U) \} \left[ \exp\{ -\lambda u_d(U) \} + \exp\left\{ \lambda \frac{e^{-\lambda u_d(U)} u_d(U)}{1 - e^{-\lambda u_d(U)}} \right\} - 2 \right].
\]

As \( u_d(U) \) is strictly monotonic in \( U \), \( Z(U) \to 0 \) as \( U \to \infty \), while for bounded \( U \), \( Z(U) = 0 \) holds if and only if the term in brackets is equal to zero, i.e., in case \( u_d \) solves

\[
\exp\{ -\lambda u_d \} + \exp\left\{ \lambda \frac{e^{-\lambda u_d} u_d}{1 - e^{-\lambda u_d}} \right\} = 2.
\]

(23)

To see that a solution to this must be unique, we have for the derivative w.r.t. \( u_d \)

\[
-\lambda \exp\{ -\lambda u_d \} - \lambda e^{-\lambda u_d} \frac{e^{-\lambda u_d} + \lambda u_d - 1}{(1 - e^{-\lambda u_d})^2} \exp\left\{ \lambda u_d \frac{e^{-\lambda u_d}}{1 - e^{-\lambda u_d}} \right\},
\]

which is clearly negative if \( v(u_d) = e^{-\lambda u_d} + \lambda u_d \geq 1 \). This follows from \( v'(u_d) = \lambda (1 - e^{-\lambda u_d}) > 0 \) together with \( v(u_d = 0) = 1 \). (Incidentally, from \( Z(U = 1/\lambda) = -I(1/\lambda) < 0 \) we can also establish that the point of intersection satisfies \( \frac{1}{\lambda} < \bar{U} < \infty \).)

**Q.E.D.**

We conclude the discussion of selective targeted disclosure with an analysis of the case where attributes have different weights:

\[
u = \alpha^1 u^1 + \alpha^2 u^2.
\]

Without loss of generality we stipulate \( \alpha^1 > \alpha^2 > 0 \). We focus on the tractable case where, as in Claim 1, \( u^1 \) and \( u^2 \) are uniformly distributed on \([\underline{u}, \overline{u}]\). When the considered sender observes the receiver’s preferences and can thus practice targeted selective disclosure, we restrict attention to the characterization of a rational expectations equilibrium where the disclosure rule is linear: \( d = 1 \) whenever \( u^1 \geq a + bu^2 \). If this rule is rationally anticipated by the receiver, then choosing \( d = 1 \) is indeed optimal if and only if

\[
\alpha^1 u^1 + \alpha^2 E \left[ u^2 | u^2 \leq \frac{u^1 - a}{b} \right] \geq \alpha^2 u^2 + \alpha^1 E \left[ u^1 | u^1 \leq a + bu^2 \right],
\]

which can be transformed to obtain

\[
u^1 \geq \frac{\alpha^1 - \alpha^2}{\alpha^1} \underline{u} + \frac{\alpha^2}{\alpha^1} \overline{u}.
\]

(24)
If (24) does not hold, \(d = 2\) is disclosed. With this rule at hand, after disclosing \(d = 1\) the expected utility equals

\[
U = \frac{3}{2} \alpha^1 u^1 - \frac{1}{2} (\alpha^1 - 2\alpha^2) u,
\]
so that \(U \in [(\alpha^1 + \alpha^2) u, \frac{3}{2} \alpha^1 u - \frac{1}{2} (\alpha^1 - 2\alpha^2) u]\). With \(d = 2\) we obtain

\[
U = \frac{3}{2} \alpha^2 u^2 + \frac{1}{2} (2\alpha^1 - \alpha^2) u,
\]
so that now \(U \in [(\alpha^1 + \alpha^2) u, \frac{3}{2} \alpha^2 u + \frac{1}{2} (2\alpha^1 - \alpha^2) u]\). Note that from \(\alpha^1 \geq \alpha^2\), which we stipulated without loss of generality, the highest value of \(U\) is attained when disclosing \(u^1 = \bar{u}, \bar{U} = \frac{3}{2} \alpha^1 \bar{u} - \frac{1}{2} (\alpha^1 - 2\alpha^2) \bar{u}\), while \(U = (\alpha^1 + \alpha^2) u\). The following characterization can now be obtained after some calculations:

\[
I(U) = \begin{cases} 
\frac{1}{\alpha^1} \alpha^1 \left( \frac{2(U-(\alpha^1+\alpha^2)u)}{3(\bar{u}-u)} \right)^2 & \text{for } U \leq U' \leq U'' \\
\frac{1}{\alpha^1} 2(U-(\alpha^1+\alpha^2)u) & \text{for } U' < U \leq \bar{U} 
\end{cases}
(25)
\]
where

\[
U' = \frac{3}{2} \alpha^2 \bar{u} + \frac{1}{2} (2\alpha^1 - \alpha^2) \bar{u},
(26)
\]
with \(U' \in (\bar{U}, \bar{U})\) for \(\alpha^1 > \alpha^2\).

We next derive \(\hat{I}(U)\) (where the receiver is not wary of the fact that the sender observes her preferences before disclosure). In this case, for the sender it is optimal to choose \(d = 1\), so as to maximize the perceived valuation, when

\[
\alpha^1 u^1 + \alpha^2 E[u^2] \geq \alpha^1 E[u^1] + \alpha^2 u^2;
\]
which transforms to

\[
u^1 \geq \frac{(\alpha^1 - \alpha^2) (\bar{u} + u)}{\alpha^1} + \frac{\alpha^2}{\alpha^1} u^2,
\]
and otherwise to disclose \(d = 2\). Again after some tedious calculations we obtain:

\[
\hat{I}(U) = \begin{cases} 
\frac{[2U-(\alpha^1+\alpha^2)\left(\frac{\bar{u}+u}{2}\right) + \frac{1}{2} (\alpha^1-\alpha^2) \left(\frac{\bar{u}+u}{2}\right)]^2}{4\alpha^1 \alpha^2 (\bar{u}-u)^2} & \text{for } \hat{U} \leq U \leq \hat{U}'' \\
\frac{2U-\alpha^2 (\bar{u}+u) - 2\alpha^1 u}{2\alpha^1 (\bar{u}-u)} & \text{for } \hat{U}'' < U \leq \bar{U} 
\end{cases}
(27)
\]
where

\[
\hat{U} = \alpha^2 u + \alpha^1 \left(\frac{\bar{u}+u}{2}\right),
(28)
\]
\[
\bar{U} = \alpha^1 \bar{u} + \alpha^2 \left(\frac{\bar{u}+u}{2}\right),
(29)
\]
\[
\hat{U}'' = \alpha^2 \bar{u} + \alpha^1 \left(\frac{\bar{u}+u}{2}\right).
(30)
\]
with $\tilde{U} \in (\tilde{U}, \tilde{u})$ for $\alpha_1 > \alpha_2$.

What now complicates the analysis is that with unequal weights $\alpha_1 \neq \alpha_2$ the sender’s strategy is no longer immediate even when he does not observe the receiver’s preferences. This is despite the fact that the sender arguably still applies the same disclosure rule to each receiver. Without loss of generality, we can limit the sender’s strategies to always disclosing the first attribute or to always disclosing the second attribute. For our subsequent derivations we need not determine which one is optimal. When the sender discloses $d = 1$, then

$$N(U) = \frac{2U - \alpha^2 (\mu + u) - 2\alpha^1 u}{2\alpha^1 (\mu - u)}$$

(31)

for $U \in \left[\alpha^1 u + \alpha^2 \frac{\mu + u}{2}, \alpha^1 \mu + \alpha^2 \frac{\mu + u}{2}\right]$. When he discloses $d = 2$, then

$$N(U) = \frac{2U - \alpha^1 (\mu + u) - 2\alpha^2 u}{2\alpha^2 (\mu - u)}$$

(32)

for $U \in \left[\alpha^2 u + \alpha^1 \frac{\mu + u}{2}, \alpha^2 \mu + \alpha^1 \frac{\mu + u}{2}\right]$.

Based on the derived expressions, one can show the following result. (The proof is available from the authors upon request.)

**Claim 4 (Appendix C).** For the case with targeted selective disclosure where $u^i$ is uniformly distributed but where the receiver applies different weights, $\alpha_1 \neq \alpha_2$, the distributions for the receiver’s perceived utility still satisfy Definition 1.