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

forthcoming *Management Science*

Florian Hoffmann† Roman Inderst‡ Marco Ottaviani§

July 2019

Abstract

This paper introduces a modeling framework to study selective disclosure of information by firms or political campaigners (senders), based on the personal information that they acquire about the preferences and orientations of consumers and voters (receivers). We derive positive and normative implications depending on: the extent of competition among senders, whether receivers are wary of senders collecting personalized data, and whether firms are able to personalize prices. We show how both senders and receivers can benefit from selective disclosure. Privacy laws requiring senders to obtain consent to acquire personal information that enables such selective disclosure increase receiver welfare if and only if there is little or asymmetric competition among senders, when receivers are unwary, and when firms can price discriminate.

Keywords: Selective disclosure, hypertargeting, limited attention, privacy regulation.

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

*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". We are grateful to Andrea Ciccarone, Marco Petterson, and especially Matteo Camboni for outstanding research assistance and thank Joshua Gans (the editor) and two anonymous referees for helpful comments. Hoffmann and Inderst gratefully acknowledge financial support from the ERC Advanced Grant 22992, Ottaviani from the ERC Advanced Grant 295835, and Inderst also from the DFG through the Research Unit "Design & Behavior" (FOR 1371).

†Erasmus University Rotterdam. E-mail: hoffmann@ese.eur.nl.

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

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

1 Introduction

Firms traditionally had two distinct ways to persuade consumers. 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 through direct marketing aimed at persuading single individuals or small groups. To implement this second strategy, firms could hire experienced salespeople to gather critical knowledge about their audiences, making it possible to tailor their messages via face-to-face contacts.

Nowadays, the greater availability of personally identifiable data on the internet blurs the distinction between these two traditional communication strategies. Sellers can systematically collect personal and detailed data about an individual's past purchasing behavior, browsing activity, or credit history, as well as the personal likes and dislikes the individual shares on social networking sites.[1] 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. On Facebook, for example, ski resorts advertise family activities to married users with kids, but stress snowboarding and party options to younger users interested in winter sports. Behavioral targeting, a.k.a. hypertargeting, along these lines combines features of remote broadcasting with features of personal selling.[2]

This paper formulates a modeling framework to study selective disclosure of information based on firms' collection of personalized data about individual consumer preferences. We realistically assume that the amount of information that can be disclosed is limited, e.g., due to time or space constraints or simply because consumers have limited attention.[3] When a firm does not know which piece of information is more likely to decrease or increase

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

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

[3] (Too much) information might "consume 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 (Dessein, Galeotti, and Santos 2016). 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.

the willingness to pay of a given consumer, the firm cannot do better but to disclose the same (restricted) information to all consumers. Only by gathering personal information about the preferences of consumers, firms are able to select among different pieces of information the one that is most likely to increase a particular consumer's willingness to choose their offering.

Specifically, our baseline setting considers a single sender attempting to persuade a single receiver to accept its offering rather than an outside option with given reservation utility. The receiver's valuation for the sender's offering is the sum of two i.i.d. components associated with two attributes of the sender's offering. The sender can disclose information that allows the receiver to learn the realization of one of the component valuations. When the sender does not know the preferences of the receiver, disclosure is non-selective in that the receiver gets to observe one randomly chosen component valuation. Instead, when knowing receiver preferences, the sender can selectively disclose the highest realization of the two component valuations.

In our main application to marketing, non-selective disclosure is coarse and simply corresponds to broadcasting of ads via billboards, newspapers or television, while selective disclosure based on firms' knowledge of individual consumer preferences captures direct marketing or online advertising targeted at the individual. The framework also applies to political campaigning, where non-selective disclosure corresponds to broadcasting of campaign messages through traditional public communication channels. Selective disclosure, instead, results when political candidates hire skilled campaigners to gather critical knowledge about individual voters' preferences and orientations.[4] In this context, we address the following questions: Does selective disclosure benefit or harm senders (firms or candidates) and receivers (consumers or voters)? When should we expect selective disclosure to arise in equilibrium? What is the role of privacy regulation?

The impact of selective disclosure on consumer welfare in equilibrium is subtle. On the one hand, consumers learn less about the bottom of the distribution since they tend to observe only the most favorable information about each firm's offering. On the other hand, consumers also implicitly learn something about the *non-disclosed* information. As we elaborate below, for a broad set of commonly used distributions of receiver valuations that satisfy logconcavity, selective disclosure increases the dispersion in the consumer's expected valuation according to a mean-preserving rotation. Ceteris paribus, selective disclosure thus ultimately is more informative and benefits consumers. When consumers are wary of selective disclosure, they tend to be worse off under selective disclosure only when there is insufficient competition and when, at the same time, firms can extract the

[4] More generally, non-selective disclosure consists of a randomized experiment, whereas selective disclosure can be seen as a manipulated experiment in which the receiver is fed with the more favorable of the two pieces of evidence.

additional valuation by *personalized pricing*.

When selective disclosure is more informative, the resulting increase in welfare can benefit both consumers *as well as* firms. However, in some important cases, firms are worse off—with a more dispersed ex-post valuation it can become less likely that the receiver chooses the sender's offering. Crucially, unless selective disclosure is (legally) not feasible or requires the consumer's consent, firms then face a commitment problem. To see this, note that, when the receiver does not expect selective disclosure, selective disclosure shifts upwards the ex-ante distribution of the consumer valuation in the sense of first-order stochastic dominance. Thus, firms would *always* want to secretly collect personal information and disclose selectively, even when they end up being damaged once consumers adjust for selectivity. Interestingly, if consumers are wary of firms' incentives and capabilities and, thus, benefit from the higher informativeness of selective disclosure, seemingly well-intended consumer-protection regulation prohibiting the collection of personal data and thereby selective disclosure could backfire. Even soft-handed regulation that requires consumers' consent for the collection of such information may help firms solve their own commitment problem vis-à-vis consumers, to the detriment of the latter.

This result ties into a lively debate about the regulation of 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 have been considering stricter regulation of consumer privacy.[5] 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.[6] The prevailing presumption—see Shapiro and Varian (1997)—is that efficiency is achieved by granting property rights over information to consumers, for example by requiring consumer consent. We show, instead, how such requirements may backfire precisely because only the collection of private information allows firms to disclose information selectively.

Further concerns, however, are raised with regards to consumers who may remain

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

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

blithely unaware of the ability of firms to collect information and communicate selectively.[7] Even though selective disclosure biases upward a receiver's perceived valuation, we show how it has no effect on the receiver's choice under symmetric competition. Ignorance can in fact be bliss when it undermines senders' commitment strategy not to collect personal information and disclose selectively.

Clearly, our model takes only a limited view of privacy. In the baseline case with fixed prices we find that there is no need to protect consumer privacy. When firms use personal information exclusively to tailor their communication by selecting which information to disclose on a personal base, enabling senders to target their disclosure strategy typically results in efficiency gains as well as in higher consumer welfare. Within our main application to marketing, however, 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. When a firm is in a monopolistic position, price discrimination based on personal information can then result in exploitative behavior, notably when consumers are unwary of firms' ability to collect and use personal data, making regulatory intervention desirable. Selective disclosure can also dampen competition by increasing perceived differentiation, from an ex-ante perspective. However, this effect is attenuated for unwary consumers, so that in this case ignorance can again be bliss. Finally, when personal information is used to price discriminate *and, at the same time,* to selectively disclose information, consumers again tend to benefit from the firms's ability to learn about their preferences when competition is sufficiently intense. In terms of policy conclusions, the strongest result delivered by our model thus is that the use of personal data to selectively disclose information increases consumer welfare when competition among senders is sufficiently intense, under all circumstances we consider.

The paper contributes to the literature on information control, disclosure, and persuasion. In our model, senders are able to control receivers' information indirectly by selectively disclosing information based on their knowledge of receiver preferences, 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 Ganuza and Penalva (2010).[8] As we stipulate that senders cannot disclose all attributes, in the spirit of Fishman and Hagerty's (1990) notion of limited attention (see also Glazer and Rubinstein (2004)), lack of disclosure does not trigger complete unraveling, thus departing from the baseline models of Grossman (1981), Milgrom (1981), and Milgrom and Roberts (1986). Our model also

[7] Tucker (2012) discusses the extent of informational asymmetry between consumers and firms in online advertising about how much personal data is being collected.

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

adds hidden information acquisition prior to the stage of selective disclosure. In addition, senders in our model cannot commit to the information structure, akin to signal-jamming à la Holmström (1999), and so can fall victim to their own incentives to secretly acquire information and disclose selectively. The non-commitment assumption also distinguishes our paper from models of *optimal* persuasion with commitment à la Rayo and Segal (2010) and Kamenica and Gentzkow (2011), where a sender commits to an information structure in an unconstrained fashion.[9] By considering unwary receivers, we also contribute to the literature on persuasion/disclosure with bounded rationality.[10]

In our model, privacy affects consumer welfare through the restrictions it imposes on the selection of disclosed information. Instead, the law and economics literature on transparency focuses mainly on the direct costs of information acquisition. The incentives to collect information may be too high when information affects the distribution of surplus (e.g., Hirshleifer 1971); for example, when information allows firms to better price discriminate.[11] To better trade off the social costs and benefits of collecting and using personally identifiable data, instead of prohibiting these practices, it has been proposed to essentially grant agents property rights over such information (e.g., Shapiro and Varian 1997). Our analysis reveals a particular twist to this policy. We show that a policy that requires consumer consent may allow firms to commit to abstain from hypertargeting, even when this would benefit consumers.

The marketing literature on targeted advertising offers a different twist on the costs of transparency. Here targeting allows firms to better restrict the scope of their marketing to those consumers who are likely to purchase in the first place (e.g., Athey and Gans

[9] Our analysis of equilibrium persuasion with multiple senders is particularly tractable given our focus on horizontal differentiation with independently distributed values. See Gentzkow and Kamenica (2017) for an analysis of optimal persuasion with multiple senders releasing information simultaneously, and Board and Lu (2018) for a related setting with consumer search and sequential information provision. 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. See also DellaVigna and Gentzkow (2010) for a survey of the literature on persuasion, including applications to marketing and political settings.

[10] See Zeckhauser and Marks (1996) for a number of insightful examples of sign posting, where senders selectively disclose information to receivers with limited attention and possibly bounded rationality. 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, instead, attend to the disclosed attribute but fail to make the appropriate inference about the undisclosed attribute, which is chosen selectively by the sender. Thus, relative comparisons across different dimensions of information play a key role in our model. 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).

[11] 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.

2010). Several recent papers (e.g., Goldfarb and Tucker 2011; Campbell, Goldfarb, and Tucker 2015, and Shen and Villas-Boas 2018) 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, overall the protection of privacy rights should always take into account the extent of competition and its benefits to consumers.[12]

The paper proceeds as follows: Section 2 sets the stage by characterizing the impact of selective disclosure in the context of a generally-applicable baseline setting, where we distinguish whether the receiver is wary or unwary. Building on this baseline analysis, our main application interprets the receiver's component valuations as horizontally differentiated match values, so that selective disclosure is based on personalized data. Sections 3 and 4 consider equilibrium disclosure in a model that allows for multiple senders who can privately choose whether to acquire information about receiver preferences before disclosing information. While Section 3 derives equilibrium for fixed prices, Section 4 allows for individualized price discrimination. For both cases, we analyze the impact of different regulatory regimes aimed at restricting senders' ability to secretly acquire information and how this impact depends on receivers' potential naïveté, the extent of competition among senders, as well as senders' potential asymmetry. Section 5 concludes and suggests avenues for future research. Appendix A reports the proofs. Appendix B collects supplementary material.

2 Selective Disclosure with One Sender and One Receiver

2.1 Baseline Setting

Our baseline model considers a single sender who aims at persuading a single receiver to accept an offering by selectively disclosing information. The starting point of our analysis is a standard and broadly applicable random-utility discrete-choice framework. The receiver has reservation utility R, which we initially take as exogenous, and utility from acceptance

$$u^1 + u^2, \tag{1}$$

[12] Another strand of the literature assumes that information disclosure is directly inconvenient for consumers, as in Casadesus-Masanell and Hervas-Drane (2015).

in the spirit of Lancaster (1966).[13] The component valuations u^1 and u^2 are identically and independently distributed with atomless distribution $F(u)$, expectation $E[u]$, and (possibly unbounded) support $(\underline{u}, \overline{u})$, with $\underline{u} + E[u] < R < \overline{u} + E[u]$.[14]

Ex ante the receiver does not know the realizations of u^1 and u^2, but can glean some information about these values on the basis of the disclosure made by the sender. Disclosure is always restricted, in the sense that the sender can disclose information that allows the receiver to directly observe *either* u^1 *or* u^2. This restriction could originate from limited attention by the receiver or, equivalently, from limitations of air time and advertising space.[15] In light of the equilibrium analysis developed in Section 3, in which each sender can decide whether to learn the realizations of u^1 and u^2 prior to disclosure, we compare the following three disclosure regimes:

(N) In the default scenario with **non-selective disclosure**, the sender discloses to the receiver a single component u^i, for $i = 1, 2$, randomly chosen independent of the realization. This regime arises naturally when the sender does not know the realizations of u^1 and u^2. The receiver thus infers that the expected valuation conditional on the observed u^i is

$$E_N\left[u^1 + u^2 | u = u^i\right] = u^i + E\left[u\right]. \tag{2}$$

(S) Under **selective disclosure**, the sender discloses to the receiver the component with the highest realization among u^1 and u^2. Selective disclosure thus requires the sender to learn the realizations of u^1 and u^2 *prior* to disclosure.[16] The fact that the sender's disclosure now depends on the realization of both component valuations u^1 and u^2, allows the receiver who correctly anticipates that disclosure is selective to infer something about the non-disclosed component. In particular, the receiver's expected valuation conditional on observing $u_d := \max\langle u^1, u^2\rangle$ is then

$$\mathcal{U}\left(u_d\right) := E_S\left[u^1 + u^2 \middle| \max\left\langle u^1, u^2\right\rangle = u_d\right] = u_d + E\left[u | u \leq u_d\right]. \tag{3}$$

[13] The equilibrium model in Section 3 endogenizes R on the basis of the best alternative offered by competing senders. The case where receivers place different weights on the two components in (1) is analyzed in Appendix B.3. When the utility components u_1 and u_2 are uniformly distributed we obtain a clean characterization. As we show, our main comparison of selective and non-selective disclosure still applies.

[14] Without loss of generality we take the support as open, so as to allow both for the possibility that it is bounded and unbounded (with either $\underline{u} = -\infty$ or/and $\overline{u} = \infty$).

[15] When a sender is informed about the values u^i a given receiver attaches to the two components of his offering, this restriction on the communication channel prevents full unravelling, thus, creating scope for disclosure strategies that depend on the receiver's preferences. The key restriction is that the sender can disclose at most one component, while disclosure of at least one component will be optimal in equilibrium due to a standard unravelling argument.

[16] The equilibrium analysis in Section 3 characterizes when the sender does choose to learn these values.

Note that regime S applies when the receiver is aware that disclosure is selective when forming expectations, which clearly is the case in equilibrium when the receiver forms rational expectations.

($\hat{S}$) **Unwary selective disclosure** arises when the sender selectively discloses the highest realization among u^1 and u^2, but the receiver wrongly anticipates non-selective disclosure, resulting in the expected valuation

$$E_{\hat{S}}\left[u^1 + u^2 \middle| \max\left\langle u^1, u^2\right\rangle = u_d\right] = u_d + E\left[u\right]. \tag{4}$$

This regime is relevant both when the unwary receiver remains unaware of the sender's incentives for selective disclosure and as a building block in the construction of equilibrium with a wary receiver, where regime $\hat{S}$ applies off the equilibrium path.

This random-utility framework of selective disclosure can be interpreted in a variety of ways. As described in the Introduction, our main application to selective disclosure based on personalized data, interprets the utility components in (1) as horizontally differentiated match values which can only be learned when knowing both the hard facts regarding the sender's offering as well as individual receiver preferences. We provide a detailed discussion of such settings in which selective disclosure requires the sender to collect information on the preferences of individual receivers in Section 3.1. However, the reduced-form random-utility model sketched above also nests settings in which component valuations u^1 and u^2 do not depend on individual receiver characteristics, such as selective disclosure of the results of product quality tests or, more generally, vertical aspects of the sender's offering.[17] Our characterization in the remainder of this section is thus relevant beyond the concrete application we consider in Section 3.[18]

2.2 Decomposition of Impact of Selective Disclosure on Receiver

To compare the receiver's expected payoff under selective and non-selective disclosure, it is useful to decompose the effect of the disclosure strategy on the expected sum of component valuations into two channels:

[17] In such a setting, non-selective disclosure could be interpreted as a randomized experiment, whereas selective disclosure would correspond to a manipulated experiment in which the receiver is fed with the more favorable of the two pieces of evidence. Suppose, for instance, that receivers are homogeneous and their willingness to pay in (1) is determined by the quality of the sender's offering, which has quality components u^i. Ex ante neither the sender nor the receiver know the values of these components, but the sender can design tests revealing the concrete realizations of u^1 and u^2, where limited attention constraints imply that the sender can disclose only one of these realizations to receivers. Non-selective disclosure would then correspond to a setting in which the sender commits to disclose the result of a particular test (e.g., the first) independent of the realization. In the selective disclosure regime, instead, the sender has discretion in choosing which of the two test results to report.

[18] Appendix A.1 collects the proofs of the results stated in the text. Appendix B.2 collects supplementary material and illustrations for parametric distributions.

- **Factual Channel**. First, the sender's disclosure strategy affects the receiver through the distribution of the disclosed variable. When the sender selectively reports $\max\langle u^1, u^2\rangle$, high realizations become more likely compared to when the sender reports a random realization u^i. This factual channel is the only one active for an unwary receiver who perceives disclosure to be non-selective such that the expected value of the undisclosed variable remains $E[u]$ even under selective disclosure.

- **Inference Channel.** When the sender's disclosure strategy is selective, the disclosed realization $u_d = \max\langle u^1, u^2\rangle$ contains indirect information about the undisclosed variable; it must be that the undisclosed realization is smaller than u_d. Only a wary receiver, who knows the sender's disclosure strategy, is able to exploit this additional information to correctly estimate the undisclosed variable as $E[u|u < u_d]$ rather than $E[u]$, adjusting for selection bias.

The impact of selective disclosure on the receiver depends on a comparison of these two channels across the disclosure regimes. To see this denote the receiver's expected payoff in disclosure regime $j = N, S, \hat{S}$ by V_j.[19] The value obtained by the wary receiver under selective disclosure can be decomposed as

$$V_S = \underbrace{V_{\hat{S}}}_{\text{Factual Channel}} + \underbrace{(V_S - V_{\hat{S}})}_{\text{Inference Channel}},$$

where $V_S - V_{\hat{S}} > 0$ is the gain from the inference channel.

Under non-selective disclosure the receiver cannot make any strategic inference, given that the sender discloses u^i independently of its actual realization. Thus, because the receiver is equally likely to observe $\max\langle u^1, u^2\rangle$ and $\min\langle u^1, u^2\rangle$, the receiver's value under non-selective disclosure is

$$V_N = \underbrace{\frac{V_{\hat{S}} + V_{\hat{T}}}{2}}_{\text{Factual Channel}} + \underbrace{0}_{\text{Inference Channel}},$$

where $V_{\hat{T}}$ is the unwary receiver's value from observing $\min\langle u^1, u^2\rangle$. Overall,

$$V_S - V_N = (V_{\hat{S}} - V_N) + (V_S - V_{\hat{S}}) = \underbrace{\frac{V_{\hat{S}} - V_{\hat{T}}}{2}}_{\Delta \text{ Factual Channel}} + \underbrace{(V_S - V_{\hat{S}})}_{\Delta \text{ Inference Channel}}. \quad (5)$$

The impact of selective disclosure on the wary receiver's value through the factual channel (first term) crucially depends on whether the unwary receiver prefers observing $\max\langle u^1, u^2\rangle$

[19] As the receiver accepts the sender's offering if and only if his expected valuation is greater than the outside option R, we obtain $V_N = \int \max\langle R, u^1 + E[u]\rangle\, dF(u^1)$, $V_S = \int \max\langle R, u_d + E[u|u \leq u_d]\rangle\, dF(u_d)^2$ and $V_{\hat{S}} = RF(R - E[u])^2 + \int_{R-E[u]}^{\infty} [u_d + E[u|u \leq u_d]]\, dF(u_d)^2$.

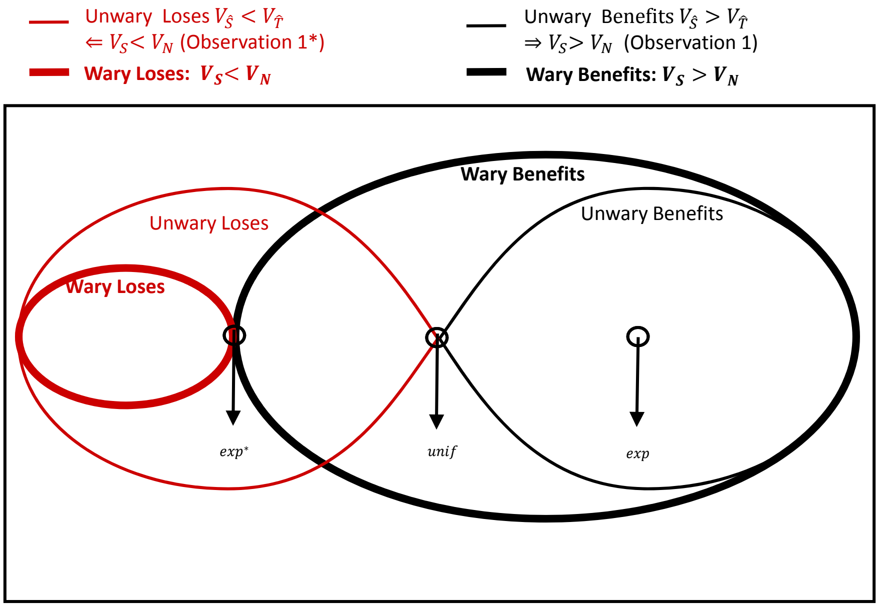

Figure 1: Illustration of Decomposition of Welfare Impact of Selective Disclosure

or $\min \langle u^1, u^2 \rangle$ at a given reservation utility R, while the impact through the inference channel (second term) is unambiguously positive. Overall, when also the factual channel is more valuable under selective disclosure, the wary receiver unambiguously benefits:

Observation 1 *A sufficient condition for the wary receiver to always (for all R) benefit from selective disclosure is that the unwary receiver's expected value is always (for all R) higher when observing* $\max \langle u^1, u^2 \rangle$ *rather than* $\min \langle u^1, u^2 \rangle$*:* $V_{\hat{S}} \geq V_{\hat{T}} \Leftrightarrow V_{\hat{S}} \geq V_N$ *is sufficient for* $V_S \geq V_N$*.*

For instance, as verified in Example B.1 in Appendix B.2, if u is exponentially distributed, the unwary receiver prefers observing $\max \langle u^1, u^2 \rangle$ rather than $\min \langle u^1, u^2 \rangle$; the wary receiver then benefits from selective disclosure. On the flip side, for the wary receiver to lose from selective disclosure it becomes *necessary*—rather than sufficient—that $V_{\hat{S}} < V_{\hat{T}}$:

Observation 1* *A necessary condition for the wary receiver's expected payoff to be always (for all R) higher under non-selective than under selective disclosure is that the unwary receiver's expected payoff is always (for all R) higher when observing* $\min \langle u^1, u^2 \rangle$ *rather than* $\max \langle u^1, u^2 \rangle$*:* $V_{\hat{S}} < V_{\hat{T}} \Leftrightarrow V_{\hat{S}} < V_N$ *is necessary for* $V_S < V_N$*.*

In general, by taking the mirror image (reflected around $u = 0$) of any distribution satisfying $V_{\hat{S}} > V_{\hat{T}}$ we obtain a distribution satisfying $V_{\hat{S}} < V_{\hat{T}}$. Thus, the two sets representing when the unwary receiver benefits or loses for all R in Figure 1 have the same size.[20] However, note the stark asymmetry between the conditions in Observation 1 and 1*, sufficient the former and necessary the latter. Thus, in Figure 1, the wary receiver benefits from selective disclosure in a superset (bold, on the right-hand side) of the set of distributions for which the unwary benefits, but loses in a subset (bold and red, on the left-hand side) of the set for which the unwary receiver loses. The underlying force that creates this asymmetry and handicaps non-selective disclosure is the positive impact of the strategic inference channel, present only under selective disclosure for the wary receiver.[21]

Impact of Selective Disclosure on Unwary Receiver. We first characterize when an unwary receiver benefits from selective disclosure, $V_{\hat{S}} \geq V_{\hat{T}}$, which, by Observation 1 is also sufficient for the wary receiver to be better off. By biasing upward an unwary receiver's perceived utility, the mistake of erroneously accepting the offering, even though $u^1 + u^2 < R$, evidently becomes larger under selective disclosure. But this is only one side of the equation. At the same time, with selective disclosure, it becomes less likely that the receiver erroneously decides against the offering, namely when actually $u^1 + u^2 > R$ holds. For symmetric distributions we obtain a clear-cut result for how the two errors trade off:

Proposition 1 *If F has a symmetric and unimodal density the unwary receiver benefits from selective disclosure, $V_{\hat{S}}(R) \geq V_N(R)$, if and only if $R \geq 2E[u]$.*

We state Proposition 1 for the natural case of unimodal densities. For distributions with U-shaped density, as an immediate corollary to this proposition, we obtain the reverse result: the unwary receiver benefits from selective disclosure, $V_{\hat{S}}(R) \geq V_N(R)$, if and only if $R \leq 2E[u]$. It then follows immediately that the unwary receiver always obtains the same expected payoff under selective and non-selective disclosure when F is uniform which, among symmetric distributions, is at the boundary between distributions with unimodal and U-shaped densities. The uniform is, thus, at the center of Figure 1, on the boundary of the set of distributions for which the unwary receiver benefits and loses. By Observation 1, we thus conclude that for uniform F, the wary receiver always (i.e., for all R) benefits from selective disclosure as $V_S - V_N = V_S - V_{\hat{S}} > 0$. As we will see in the next section, the wary receiver benefits from selective disclosure much more generally.

To get some intuition for the result in Proposition 1 consider the Cartesian product of the supports of u^1 and u^2 as depicted in the two panels of Figure 2 which correspond to

[20] Observation 1 and 1* of course hold equally also for fixed R. Thus distributions for which the value comparison depends on R are outside of the receiver benefits/loses regions in Figure 1, cf. Figure 5 in Appendix B.2.

[21] By (5) we have $V_{\hat{S}} > V_{\hat{T}} \Rightarrow V_S > V_{\hat{T}} \Rightarrow V_S > V_N$ and Observation 1's sufficient condition can be replaced by the less stringent $V_S > V_{\hat{T}}$. Similarly, from $V_{\hat{S}} < V_{\hat{T}} \Leftarrow V_S < V_{\hat{T}} \Leftarrow V_S < V_N$, Observation 1*'s necessary condition can be replaced by the more stringent $V_S < V_{\hat{T}}$. This additional asymmetry further highlights that there are many more distributions for which selective disclosure benefits rather than harms the wary receiver.

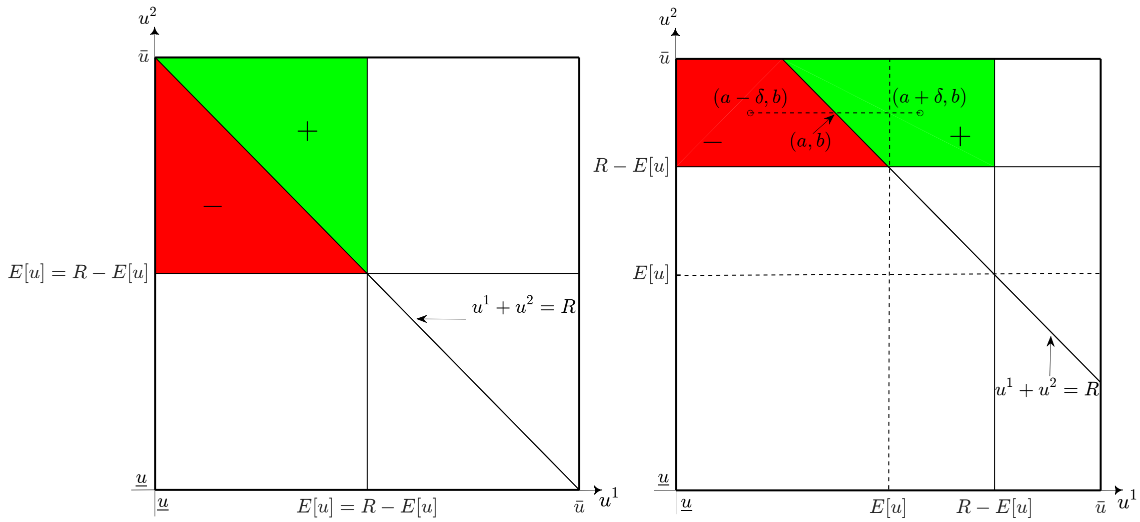

Figure 2: Welfare impact on unwary receiver.

two different values of the outside option R. With non-selective disclosure the receiver who observes u^1 is indifferent between accepting and rejecting the offering at $u^1 = R - E[u]$. Under selective disclosure, the unwary receiver still accepts when observing a realization above $R - E[u]$ as under non-selective disclosure, but now accepts not only when $u^1 \geq R - E[u]$ but also when $u^2 \geq R - E[u]$. Thus, all points (u^1, u^2) in the shaded region $[\underline{u}, R - E[u]] \times [R - E[u], \overline{u}]$ in the figure correspond to additional acceptances by the unwary receiver induced by the selectivity of disclosure.

Within the shaded region, the receiver loses when $u^1 + u^2 < R$, in (red) subregions to the left of the iso-payoff corresponding to R (increase in false positives, marked with a $-$ sign), and gains in (green) subregions to the right (increase in correct positives, marked with a $+$ sign). As is easily seen from both panels of the figure, for symmetric distributions, the total area of gains ($+$) and losses ($-$) is the same. The net effect then depends on the relative probability of points corresponding to comparable gains and losses.

Consider now, first, the left-hand panel of Figure 2 depicting the case where $R = 2E[u]$ such that the receiver is indifferent at the prior. As is immediate from the graph, given symmetry of F, in this case, points corresponding to gains and losses of the same size have the same probability. Hence, the expected payoff of the unwary receiver is unaffected by selective disclosure, $V_{\hat{S}}(R) = V_N(R)$; this indifference between the offering and the outside option at the prior can be seen as a condition of equipoise.

Consider next, the right-hand panel of Figure 2 depicting a case where $R \geq 2E[u]$. Due to symmetry of F, the total area of gains ($+$) and losses ($-$) is again the same. Further, as illustrated in the figure, for each point in the $+$ region with a given gain δ we can again find a corresponding point in the $-$ region with a loss of same size. However, the relative probability of these two points now no longer needs to be the same. In particular, the points corresponding to gains to the right of the iso-payoff of level R are closer to the expectation $E[u]$ than the points corresponding to similar losses to the left of the iso-payoff. Thus, if the distribution F is unimodal, gains (generated by the increase in correct positive decisions) weigh more than losses (generated by the increase in false positive decisions) and the unwary receiver benefits from selective disclosure.

When the distribution F is uniform, all points in Figure 2 are equally likely, so that for all reservation utilities R the increase in expected payoff associated to the additional correct positives is exactly offset by the increase in false positives, so that the unwary receiver always obtains the same expected payoff under selective and non-selective disclosure.

2.3 Impact of Disclosure on Distributions of Expected Valuation

To characterize the impact of selective disclosure on a wary receiver we now turn to a more detailed analysis of the distributions of expected valuations induced by different disclosure regimes. Given the information disclosed by the sender and the conjectured disclosure strategy, the receiver updates the valuation for the sender's offering before deciding whether to accept or reject, according to (2), (3), and (4). Next, we derive the ex ante distribution of the receiver's expected valuation resulting in the different disclosure regimes:

Non-Selective Disclosure. Given that the sender discloses a single variable, u^i, independently of its realization, the distribution N of the expected valuation $u^i + E[u]$ for the receiver conditional on observing a single variable u^i is

$$N(U) = F(U - E[u]) \text{ with } U \in (\underline{u} + E[u], \overline{u} + E[u])$$

for all possible realizations of the expected sum U. For example, for a uniform $F(u^i) = u^i$ on $(0,1)$, we have $N(U) = U - 1/2$ with support $(1/2, 3/2)$, as depicted in Figure 3.

Selective Disclosure. Rewriting expression (3), the receiver who is aware that the sender discloses $\max\langle u^1, u^2\rangle$ infers that the expected sum conditional on the observed realization $u_d = \max\langle u^1, u^2\rangle$ is

$$\mathcal{U}(u_d) = u_d + \left(\int_{\underline{u}}^{u_d} u \frac{f(u)}{F(u_d)} du\right) = u_d + \left(u_d - \int_{\underline{u}}^{u_d} \frac{F(u)}{F(u_d)} du\right) = 2u_d - \frac{L(u_d)}{F(u_d)}, \quad (6)$$

where we used integration by parts and the definition $L(u) := \int_{\underline{u}}^{u} F(y)\,dy$ of the left-hand integral of the distribution function. The term $L(u_d)/F(u_d)$ is equal to the *mean-advantage-over-inferiors* function, as defined by Bagnoli and Bergstrom (2005, page 249).

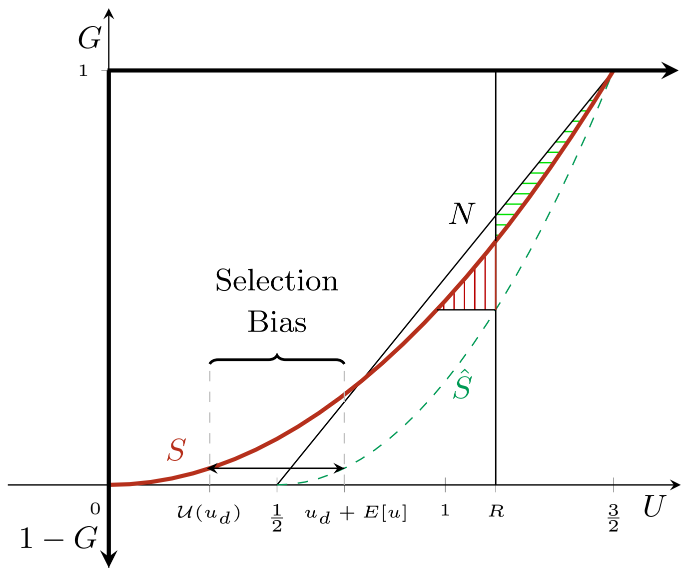

Figure 3: Comparison of distributions of posterior evaluations $G = N, S, \hat{S}$ for the uniform distribution, satisfying Property 2.

The distribution S of the expected sum $\max\langle u^1, u^2\rangle + E[u|u \leq \max\langle u^1, u^2\rangle]$ conditional on observing $\max\langle u^1, u^2\rangle$ then is

$$S(U) = F\left(\mathcal{U}^{-1}(U)\right)^2 \text{ with } U \in (2\underline{u}, \overline{u} + E[u]),$$

where the inverse function $\mathcal{U}^{-1}$ is well defined given that $\mathcal{U}$ is monotone increasing. In the uniform example, we have $E[u|u \leq u_d] = u_d/2$ so that $\mathcal{U}(u) = 3u_d/2$. Thus, the distribution of the expected valuation is $S(U) = (2U/3)^2$ with support $(0, 3/2)$, bold in Figure 3.

Unwary Selective Disclosure. In this case, the receiver is overoptimistic about the undisclosed variable and fails to adjust downward the expectation to account for selection bias

$$\delta(u_d) = \underbrace{E[u] - E[u|u \leq u_d]}_{\text{Selection Bias Adjustment}} > 0, \tag{7}$$

the difference between (3) and (4). The distribution $\hat{S}$ of the perceived valuation U, given the receiver's wrong beliefs about the disclosure regime, then is

$$\hat{S}(U) = F(U - E[u])^2 \text{ with } U \in [\underline{u} + E[u], \overline{u} + E[u]].$$

In the uniform example, the distribution of the expected perceived valuation by the unwary receiver is $\hat{S}(U) = (U - 1/2)^2$ with support $(1/2, 3/2)$, dashed in Figure 3.

Impact on Unwary Receiver. Figure 3 illustrates the following relation between the distribution N of expected valuation under non-selective disclosure and the distribution $\hat{S}$ of the expected valuation under selective disclosure as perceived by the unwary receiver:

Property 1 (FOSD) *Distribution $\hat{S}$ first-order stochastically dominates distribution N,*

$$\hat{S}(U) < N(U) \text{ for } U \in (\underline{u} + E[u], \overline{u} + E[u]).$$

This FOSD property is completely general and is based on the stochastic dominance property of order statistics: $F^2(u) < F(u)$. The FOSD property will be a key determinant of the sender's incentives for choosing the disclosure regime in the equilibrium analysis of the full model presented in Section 3. Holding fixed the beliefs of the receiver, a move to selective disclosure results in an increase of the probability that the offering is accepted for any given reservation utility R.

Impact on Wary Receiver. For the uniform distribution in Figure 3 the following rotation property holds:

Property 2 (Clockwise Rotation) *Distribution S is a mean-preserving clockwise rotation of distribution N, meaning that S crosses N once and from above at one $\tilde{U}$ in the interior of the support*

$$S(U) \gtreqless N(U) \text{ for } U \lesseqgtr \tilde{U}.$$

The (strict) Clockwise Rotation property has two important welfare implications. First, combined with the preservation of the mean that follows from the law of iterated expectations, Clockwise Rotation directly implies that S is a mean-preserving spread of N. Thus, we must have

$$V_S = \int \max\langle U, R\rangle \, dS(U) > \int \max\langle U, R\rangle \, dN(U) = V_N$$

for all choices of R because $\max\langle U, R\rangle$ is a convex function. Hence, the wary receiver unambiguously benefits from selective disclosure.[22] In Figure 3 the shaded area between the two distributions and to the right of the reservation utility (drawn as a vertical line at R) represents the difference in the receiver's value.[23] Second, clockwise rotation also

[22] Clockwise Rotation is a strengthening of the convex order—see e.g. Shaked and Shantikumar (2007, Theorem 3.A.44 on page 133)—and, thus, sufficient for the wary receiver to benefit from selective disclosure.

[23] For an economic interpretation, turn around axes and place the origin at the top-left corner. The reliability function $1 - G(U)$ can then be seen as the quantity demanded when the marginal consumer is offered utility U. Thus, the expected consumer surplus at R is equal to the area below the demand (i.e., to left of the distribution) and above the reservation utility R (to the right of R with the initial origin).

determines whether the sender prefers selective disclosure when facing a wary receiver, which is the case if and only if $R > \tilde{U}$.[24]

In Appendix B.2, we provide a characterization of the set of distributions for which S is a clockwise rotation of N (Proposition B.1) as well as sufficient conditions (Lemma B.1). There we also verify that all the random variables with logconcave F listed in Bagnoli and Bergstrom (2005), such as power distributions (including uniform), gamma(α, β) as well as Weibull with shape $\alpha \leq 1$ (including exponential), extreme value Gumbel, Pareto with $\alpha > 1$ (for which the expectation exists), normal, lognormal, Fisher-Snedecor F, and beta(α, β) with $\beta \geq 1$ satisfy the clockwise rotation property. In fact, whenever the distributions S and N cross only once, the rotation is clockwise if and only if the left-hand integral $L(u) = \int_{\underline{u}}^{u} F(y)\, dy$ is logconcave at the upper bound, a property that is implied by logconcavity of the distribution F, in turn implied by logconcavity of the density f.[25] In the boundary case with loglinear L (corresponding to F positive exponential, the mirror image of the negative exponential), selective and non-selective disclosure induce identical distributions $S = N$ (see Example B.6 in Appendix B.2). As represented in Figure 1, the positive exponential distribution thus is the boundary case between the sets of distributions for which the wary receiver always benefits and always loses from selective disclosure. It is then easy to see that logconvexity of L at the upper bound of the support implies that S crosses N from below at the last crossing (see Lemma B.2 in Appendix B.2) and thus that the receiver is harmed by selective disclosure when R is sufficiently high.[26]

The discussion above illustrates the subtleties of the impact of selective disclosure on the wary receiver's welfare. Still, as shown in Section 2.2, the direct benefit of the strategic inference channel tilts the welfare comparison in favor of selective disclosure (cf. Figure 1). In particular, the wary receiver always benefits from selective disclosure for the large class of commonly used log-concave distributions, for which the Clockwise Rotation property holds. Hence, for the remainder of the paper we assume that Property 2 (Clockwise Rotation) is satisfied, which in turn is sufficient (but not necessary, cf. Example B.11 in Appendix B.2) for the wary receiver to benefit from selective disclosure.

[24] This is easily seen by noting that the probability of acceptance is higher under selective disclosure, $1 - S(R) > 1 - N(R)$, when the outside option is high, $R > \tilde{U}$ (*niche market*), whereas it is higher under non-selective disclosure, $1 - S(R) < 1 - N(R)$, when the outside option is low, $R < \tilde{U}$ (*mass market*).

[25] These results follow from Prékopa's (1973) Theorem, which guarantees that logconcavity is preserved by integration; see, for example, An (1998) and Bagnoli and Bergstrom (2005).

[26] Appendix B.2 displays distributions with logconvex L (i.e., with bottom tail thicker than the positive exponential distribution) for which the wary receiver always loses from selective disclosure. In particular, there we also provide formal statements characterizing when the wary receiver prefers non-selective rather than selective disclosure.

2.4 Selective Non-Disclosure

The FOSD and Clockwise Rotation properties—on which the results of our analysis in Section 3 hinge—hold well beyond the specific model of selective disclosure we considered above. To illustrate the broader applicability of our results, we now sketch an alternative model of selective *non*-disclosure in which these properties also hold (see Appendix B.4 for the complete analysis).[27] The setting features a sender who has to choose between disclosing and not-disclosing a uni-dimensional signal, rather than having to choose between which of two signals to disclose. Concretely, the sender obtains with some probability the information that allows the sender to perfectly learn the receiver's valuation of the sender's offering. An informed sender then decides whether or not to allow also the receiver to learn this valuation by disclosing the relevant information. Else, when the sender remains uninformed, he cannot disclose anything (or optimally chooses not to) as in the well-known model of Dye (1985), Farrell (1986), Jung and Kwon (1998), and Shavell (1994). A sender who learns that the receiver has a low valuation then pools with the sender who does not become informed, preventing full unravelling and providing scope for selective *non*-disclosure.

Recall that in our baseline model of Section 2.1 selective disclosure required the sender to know receiver valuations, whereas non-selective disclosure did not rely on such information. In the current setting, (more) selective disclosure based on better information now similarly corresponds to a situation in which the sender is informed with higher probability. As we show in Appendix B.4 such an increase in the probability with which the sender is informed, leads to (i) a mean-preserving Clockwise Rotation in the ex-ante distribution of the receiver's perceived utility when anticipated (Property 2), and (ii) a FOSD shift in the distribution when the switch is not anticipated (Property 1).[28] Given that most of the analysis reported below only relies on these two properties, the results are valid also for this model.

Building on the core characterization of selective disclosure based on Properties 1 and 2, the paper proceeds to derive the equilibrium disclosure strategy when multiple senders with horizontally differentiated offerings compete for receivers who make individual (purchase) decisions. In the context of our main application to marketing, we also analyze the impact of privacy regulation limiting firms' scope for selective disclosure, depending on the wariness of consumers, the extent of competition among possibly asymmetric firms,

[27] In fact, within the concrete selective non-disclosure setting outlined in Appendix B.4, Properties 1 and 2 hold generally, i.e., independent of the underlying distribution of receiver preferences.

[28] In order to make precise how Properties 1 and 2 extend to this setting, denote the ex-ante distribution of the receiver's perceived valuation when the receiver correctly believes that the likelihood of facing an informed sender is low by $N(U)$. Similarly, let $S(U)$ and $\hat{S}(U)$ denote the respective distribution when the sender is informed with high probability and this is anticipated or, respectively, not anticipated by the receiver.

and their ability to practice individualized price discrimination.

3 Equilibrium Selective Disclosure

3.1 Interpretation in Terms of Horizontal Match Information

The baseline random utility model introduced in Section 2.1 takes as primitives the receiver's component valuations u^i, for $i = 1, 2$. While this reduced-form specification captures a variety of settings, in our main application analyzed in this section these component valuations represent horizontally differentiated match values. In particular, for a given offering, the ranking of component valuations depends on the personal characteristics of the respective receiver, so that only by collecting *personalized* data are senders able to tailor their disclosure to *individual* receivers' preferences. To fix ideas, we now sketch one particular microfoundation along these lines.

In this microfoundation, ex-ante the sender has perfect and private information about the hard facts relevant for the offering (e.g., product features, specification of product attributes, details of political campaign, etc.), while receivers have perfect and private information about their personal preferences. However, *match values* u^1 and u^2 for a particular receiver can only be learnt by combining knowledge of the hard facts about the offering with the preferences of an individual receiver. One way to formalize the mapping of hard facts about the sender's offering and individual receiver preferences into utility u^i is to consider a function $u^i := u(x^i, y^i) \in \mathbb{R}$, where $x^i \in X$ captures facts and $y^i \in Y$ captures receiver preferences. This gives rise to the above random-utility reduced-form model whenever knowledge of *either* $x = (x^1, x^2)$ or $y = (y^1, y^2)$ results in the same conditional distribution over u^i, i.e., $F(u^i|x) = F(u^i|y) = F(u^i)$, so that ex-ante differentiation is purely horizontal.[29] Next, we sketch a concrete setting for which this is the case.

Suppose that the receiver has private information about his personal preferences defined over a set of possible features, but does not know which of these features the sender's offering actually entails. Only knowledge of both the relevant features (the hard facts about the sender's offering) and receiver preferences allows learning of the receiver's willingness to pay for the sender's offering. Concretely, suppose there is a set of possible features indexed by $j \in J = [0, 1]$. A receiver's willingness to pay, y_j, for a particular feature is an i.i.d. draw from an atomless distribution $F(y_j)$. A particular sender's offering entails only *two* of these features which we denote by x^i, $i = 1, 2$, and which correspond to two i.i.d.

[29] In particular, for a sender who has not acquired information about receiver preferences, the likelihood with which a particular match value is realized is independent of the sender's "type" (the hard facts about the offering).

draws from a standard uniform distribution on J.[30] Thus, the receiver's willingness to pay for the sender's offering is given by (1), with $u^i = y_{x^i}$, $i = 1, 2$, distributed as $F(u^i)$.[31,32]

In settings with horizontally differentiated match values, the mapping of hard facts about the sender's offering into a receiver's match utility depends on individual receiver characteristics. In particular, any specific component of the sender's offering might appeal to one receiver but not to another. Hence, disclosing a given piece of such horizontal match information (in the form of either x^1 or x^2) can lead to a higher or lower valuation at a particular receiver where the direction is ex-ante uncertain. This means that implementing the selective disclosure strategy with more than one receiver requires the sender to collect personalized information on each individual receiver. An implication of selective disclosure in these settings is that the disclosed horizontal match information, and, thus, the valuation component u^i a particular receiver can learn, differs across receivers, hence, requiring a private communication channel. In contrast, non-selective disclosure can be implemented by disclosing the same (randomly chosen) piece of information to all receivers via public communication.

Building on this setting with horizontally differentiated match values and selective disclosure based on the acquisition of personalized data, we next turn to an equilibrium model with multiple senders.

3.2 Model with Multiple Senders

Suppose now that the choice of disclosure rule is made strategically by (multiple) senders. In particular, let $M \geq 2$ be the set, as well as the number, of the alternatives from which the receiver can choose one. For each offering, we denote the component valuations by u_m^i for $i = 1, 2$, whose sum gives the receiver's utility from acceptance as in (1). Receiver preferences are independent across senders and we allow for heterogeneous distributions $F_m(u_m^i)$ with support $(\underline{u}_m, \overline{u}_m)$. To ensure that the different offerings are indeed in competition we stipulate that the valuations are not too different, so that for any pair (m, m') it holds that

$$\overline{u}_m + E[u_m] > \underline{u}_{m'} + E[u_{m'}], \tag{8}$$

where we dropped the superscripts i denoting the attributes.[33]

[30] Without knowledge of the x^i, $i = 1, 2$, the receiver's willingness to pay is a random draw from $y^i := (y_j)_{j \in J}$ for $i = 1, 2$, given that the receiver does not know which two of the features $j \in J$ are relevant.

[31] Formally, in this case we have $u(x^i, y^i) = \int_{j \in J} y_j \delta\left(j - x^i\right) dj = y_{x^i}$, where δ is the Dirac delta function.

[32] Appendix B.1 presents an alternative foundation à la Salop, in which match values are decreasing in the distance between actual product attributes and the specification of attributes that is most preferred by consumers.

[33] Recall that distribution S has support $(2\underline{u}, \overline{u}+E\left[u\right])$, while the distribution of N is $(\underline{u}+E\left[u\right], \overline{u}+E\left[u\right])$. Condition (8) thus is necessary and sufficient for the supports of any combination of $G_m = S_m, N_m$ and

Unless we explicitly state otherwise, we focus on the case in which all alternatives are offered by strategic senders, so that M also corresponds to the set of senders. Still, the framework allows for the possibility that a subset of alternatives comprises outside options, for which the perceived value is not affected by senders' strategies. In particular, we will repeatedly consider the case of a single (monopolistic) sender, where the receiver's alternative is an outside option of fixed value R as in the foundational analysis above.[34]

With respect to senders' preferences, we only need to specify that each sender m is strictly better off when the receiver chooses his option m.[35] Senders thus choose their disclosure strategy—selective or non-selective disclosure—in order to persuade the receiver to choose their offering. For a sender to use selective disclosure, he must have learned the receiver's preferences. In this section we first suppose that all senders can do so without (regulatory) restrictions and denote the strategies by $s_m \in \{y, n\}$, so that y (yes) means that sender m learns the receiver's preferences and n (no) corresponds to not acquiring the information.

We thus consider now the following game:

- At $t = 1$, each sender chooses whether or not to acquire information, $s_m \in \{y, n\}$. The baseline assumption in Section 3.3 is that information acquisition is an unobservable hidden action, while in Sections 3.4 and 3.5 we discuss how regulation may imply observability (and thereby commitment).

- At $t = 2$, each sender discloses to the receiver information in the form of hard facts about one component of his offering thereby revealing to the receiver either u_m^1 or u_m^2. This induces an updated perceived valuation U_m, which clearly depends both on what the receiver observed as well as on his belief about the sender's disclosure strategy.[36] When senders are firms selling a product or service they may, at this stage, also engage in personalized pricing as analyzed in Section 4.

- Finally, at $t = 3$, the receiver chooses offering $m \in M$ that has the highest perceived utility, U_m, and randomizes with equal probability in case of a tie.

$G_{m'} = S_{m'}, N_{m'}$ to strictly overlap.

[34] Such a fixed outside option can be easily accomodated by specifying that $u_m^i = R/2$ with probability one.

[35] Given that the distributions of preferences across alternatives are independent, we in particular need not specify whether sender m receives a different (or the same) payoff when alternatives $m' \neq m$ or $m'' \neq m$ are chosen.

[36] As is easy to show, the restriction to selective and non-selective disclosure is in fact without loss of generality if we restrict attention to monotone equilibria.

3.3 Non-Commitment Regime

Recall that, for our baseline model without regulation, we assume that the choice of s_m is an unobservable hidden action and that there is also no other way for the senders to commit to a certain information acquisition or disclosure strategy. Denote then the receiver's belief about sender m's information acquisition strategy for the case of pure strategies by $\widehat{s}_m \in \{y, n\}$. We solve the game backwards. At $t = 2$, a sender who chose $s_m = n$ can only disclose non-selectively. When this was anticipated by the receiver, $\widehat{s}_m = n$, recall that the distribution of U_m is given by $G_m(U_m) = N_m(U_m)$. Suppose now that $s_m = y$. Given that receivers place the same weight on both attributes and given that the fit for each attribute i is distributed according to the same distribution function $F_m(u^i_m)$ it is then easy to show that such an informed sender will disclose selectively, independently of receivers' beliefs $\widehat{s}_m$.[37] Hence, a receiver with belief $\widehat{s}_m = y$ also anticipates selective disclosure, such that, when a receiver rightly anticipates the sender's acquisition of information, $\widehat{s}_m = s_m = y$, then $G_m(U_m) = S_m(U_m)$, while when the receiver holds the wrong beliefs $\widehat{s}_m = n$, though it holds that $s_m = y$, we have $G_m(U_m) = \hat{S}_m(U_m)$.

We can then formalize the senders' incentives to become better informed in order to disclose selectively as follows. For a given realization of U_m, $w_m(U_m) = \prod_{m' \in M \setminus m} G_{m'}(U_m)$ is the "winning" likelihood with which the receiver chooses alternative m.[38] Hence, from an ex-ante perspective, alternative m is chosen with probability

$$q_m = \int w_m(U_m) dG_m(U_m).$$

The following is now an immediate implication of the fact that $G_m(U_m) = \hat{S}_m(U_m)$ dominates $G_m(U_m) = N_m(U_m)$ in the strict FOSD order and that $w_m(U_m)$ is non-decreasing.

Lemma 1 *Suppose senders are unable to commit to their information acquisition strategy. Then, in the unique equilibrium all senders choose to acquire information about the receiver's preferences and then disclose selectively ($s_m = y$ for all $m \in M$).*

We showed already that when the receiver compares this offering to an exogenous reservation value, the receiver benefits from such selective disclosure. This property now extends to sender competition. To formalize this insight, note first that the receiver realizes $U^{(1)} = \max_{m \in M} U_m$. Note also that in equilibrium the receiver's expectations about senders' strategies hold true ($\widehat{s}_m = s_m$). Pick now some $m \in M$ and denote by $U^{(1:M \setminus m)}$ the maximum over all remaining senders, with distribution $G^{(1:M \setminus m)}(U^{(1:M \setminus m)})$. For given

[37] See however Appendix B.3 for the case with asymmetric weights.

[38] We use here that, for all possibilities, $G_m(U_m)$ does not have mass points, though the subsequent analysis can be readily extended to the case where distributions have atoms as would be the case, e.g., in the alternative selective non-disclosure setting outlined in Appendix B.4.

U_m, the receiver accordingly realizes $E\left[\max\left\{U^{(1:M\setminus m)}, U_m\right\}\right]$. Given that the expression in brackets is a convex function of U_m, reflecting the fact that taking m is an *option* for the receiver, the receiver obtains a higher expected utility after a mean-preserving spread in $G_m(U_m)$ (which is implied by the rotation property).

Lemma 2 *A switch from non-selective to selective disclosure by any sender m for which Property 2 holds strictly benefits a receiver who is aware of this, regardless of other senders' strategies.*

From now on, we assume that Property 2 holds for every sender m, so that S_m is a clockwise rotation of N_m. Combining Lemmas 1 and 2, we have the following equilibrium characterization:

Proposition 2 *When senders are unable to commit to an information acquisition strategy, the unique equilibrium outcome maximizes receiver's utility through selective disclosure by all senders.*

3.4 Equilibrium with Commitment

We now turn to circumstances in which senders have the ability to credibly commit to a certain information acquisition strategy.[39] This is so, in particular, when information acquisition requires either the direct cooperation or at least the consent of receivers, in which case receivers in fact directly observe a sender's attempt to become better informed about receiver preferences in order to disclose more selectively. We relate this to the question of (optimal) privacy regulation after the analysis of the equilibrium outcome. Note that large companies might also be able to commit not to collect information through reputation. In this regard see the discussion in Agrawal, Gans, and Goldfarb (2018, pages 189-190) about Apple's privacy strategy.

In contrast to the preceding analysis without commitment, now a given sender m's preferred choice depends crucially on the distribution of a receiver's next best alternative, which we denoted by $G^{(1:M\setminus m)}(U^{(1:M\setminus m)})$. To illustrate, we return to the simple case analyzed in the preceding section, where the receiver's best alternative has a deterministic value (that is, $U^{(1:M\setminus m)} = R$). The receiver would then choose sender m's preferred alternative with probability $1 - S(R)$ when the sender discloses more selectively, rather than with probability $1 - N(R)$. The probability that the receiver accepts the sender's offering is strictly higher under more selective disclosure if and only if R lies to the right of the intersection of $S(U_m)$ and $N(U_m)$, while otherwise it is strictly lower. That is, in this

[39] While we do not assume that senders can commit to a particular disclosure strategy, since information acquisition is a prerequisite for any disclosure strategy that conditions on the realization of u^i, $i \in \{1, 2\}$, commitment not to acquire information in fact also implies commitment not to disclose selectively.

example the sender prefers information acquisition and subsequent selective disclosure if the receiver's preferred alternative is sufficiently attractive (high R), while he otherwise prefers not to acquire information, which then essentially commits the sender to disclose non-selectively. We now generalize this insight to the case where $U^{(1:M\setminus m)}$ is stochastic. For this we use the following auxiliary result:

Lemma 3 *Consider a single sender m and suppose the distribution of the receiver's next best alternative, $U^{(1:M\setminus m)}$, undergoes a shift resulting in a distribution that dominates in the likelihood ratio order. Then, if the sender weakly prefers selective disclosure before the shift, where the choice is observed by the receiver, the preference becomes strict after the shift. Likewise, if the sender weakly prefers non-selective disclosure after the shift, the preference must be strict before the shift.*

To use this result, we need to map the comparative analysis into the model's primitives. Here, we focus on a comparative analysis in terms of competition, as expressed by the number of senders M.[40] To do so in a compact way, so that M is the only relevant variable to consider, we suppose that all senders $m \in M$ are symmetric, i.e., that the distributions $N_m(U_m)$ and $S_m(U_m)$ are identical across senders. Then pick any sender m and suppose that all other senders $m' \neq m$ disclose selectively such that $G_{m'}(U_{m'}) = S(U_{m'})$ (where we dropped the subscript due to symmetry). The receiver's best alternative (to m) is then distributed according to $G^{(1:M\setminus m)}(U_m) = S^{M-1}(U_m)$. Thus, as the number of senders increases from M to $M+1$ we obtain for the likelihood ratio of the receiver's outside option that

$$\frac{g^{(1:(M+1)\setminus m)}(U_m)}{g^{(1:M\setminus m)}(U_m)} = \frac{M}{(M-1)} S(U_m), \tag{9}$$

which is increasing in U_m. By property (9) we can now invoke Lemma 3 to obtain the following result:

Proposition 3 *Suppose that senders are symmetric and can commit to a information acquisition strategy (e.g., as their information acquisition strategy s_m is observed by the receiver). Then there is a finite threshold M' such that for all $M \geq M'$ there exists a unique equilibrium where all senders disclose selectively ($s_m = y$ for all $m \in M$), while for $M < M'$ this outcome is not an equilibrium.*

[40] An alternative way would be to hold M fixed and to consider a switch in the respective distribution $G_{m'}$ of any other sender. To illustrate, suppose $M = 2$ and that u_1^i is distributed uniformly over $[0, 1]$. For sender $m = 2$ take $G_2(U_2) = N_2(U_2)$ and suppose that u_2^i is also distributed uniformly but with support $[0, \overline{u}_2]$, for $0 < \overline{u}_2 < 3$. Then $m = 1$ prefers to become informed and disclose selectively, resulting in $G_1(U_1) = S_1(U_1)$, if $\overline{u}_2 > 3/4$ (i.e., if the receiver's alternative is relatively attractive), but prefers not to become informed if $\overline{u}_2 < 3/4$ (i.e., if the receiver's alternative is relatively unattractive).

Proposition 3, thus, provides an equilibrium characterization for (sufficiently) large M. Intuitively, as the number of senders increases, it becomes more and more likely that the receiver obtains a highly valuable offer elsewhere, so that from the perspective of sender m, more and more probability mass shifts to the upper tail in the distribution of the receiver's best "outside option", $U^{(1:M\backslash m)}$. As a consequence, each sender m prefers to disclose selectively in order to increase the probability of particularly high realizations of U_m.[41] Proposition 3 also covers the converse case: When M falls below the threshold M', there no longer is an equilibrium where all senders choose to disclose selectively.[42] There, the same intuition applies as in case of a deterministic outside option $U^{(1:M\backslash m)} = R$ with a low value of R: A sender does not want to disclosure selectively when, given a low M, the receiver's alternative is likely to be relatively unattractive. Taken together, receiver welfare is, thus, maximized if and only if competition among senders is sufficiently strong.

3.5 Implications for Regulation

We next evaluate the welfare implications of various commonly observed regulatory tools that put restrictions on senders' (firms') efforts to collect data about individual receivers (consumers). Initially, we conduct this welfare analysis within our baseline model in which personalized data is just used to selectively disclose information to wary receivers. We subsequently extend the analysis allowing both for unwary receivers and for price discrimination based on firms' knowledge of consumer preferences.

Absent regulation, receivers may not be able to control the extent to which firms collect personal data, for example through past purchases that reflect also on consumer preferences for the current offering. Within our baseline setting, Proposition 2 implies that all senders would indeed collect such information in order to disclose selectively thereby maximizing receiver welfare. Hence, consumer regulation that entails an outright prohibition of practices through which firms collect personalized data would hurt consumers by Lemma 2. 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 consumer preferences and disclose selectively, but cannot do so because consumers do not observe whether firms collect and use personalized data.

Regulation that prescribes consumer consent can provide exactly this commitment, which from Proposition 3 is in the interest of firms, at least when M is low, but not

[41] This result can easily be generalized to the case of asymmetric senders, as long as distributions N_m and S_m have the same support across senders and sender types, as characterized by these distributions, are drawn from a finite set.

[42] We do not provide a complete characterization of equilibrium for the case with low M, in which existence can only be guaranteed by allowing for mixed strategies.

in the interest of consumers (see Lemma 2). We conclude that regulation can backfire and decrease consumer welfare even if it just requires consumer permission, rather than prohibiting the collection and use of personal data. In the context of our model when the Clockwise Rotation property holds, consumer protection does not justify regulatory intervention in firms' data collection activities when personalized data is just used to selectively disclose information to wary consumers. This conclusion no longer holds when consumers are unwary, as we show next, and when firms can further use their knowledge of consumer preferences to price discriminate, as we show in Section 4.

Unwary Receivers. Given the novelty of hypertargeting technology, receivers may well be unaware of senders' capability to collect and use data for selective disclosure. Suppose, thus, that receivers are unwary in that they believe that $\widehat{s}_m = n$ irrespective of the senders' incentives to acquire information and disclose selectively.[43] Then, when a sender chooses $s_m = y$ instead of $s_m = n$, the distribution of perceived valuation of an unwary receiver shifts in the FOSD order from $N_m(U_m)$ to $\hat{S}_m(U_m)$.[44] When there are no (regulatory) restrictions this immediately implies that senders will always choose selective disclosure if receivers are unwary of this (cf. Lemma 1). Our subsequent analysis thus focuses exclusively on receiver welfare.

In equilibrium, when an unwary receiver observes u^i_m, his perceived value U_m is inflated and exceeds the *true* conditional expected value by $E[u^i_m] - E[u^i_m \mid u^i_m \leq u_{d_m}]$, where $u_{d_m} = \max \langle u^1_m, u^2_m \rangle$. In Section 2.2 we already remarked that this shuts down one channel through which wary receivers can benefit from selective disclosure (strategic inference). Interestingly, with a monopolistic sender (and thus an outside option of fixed value for the receiver) we also have shown there that, despite his inflated expectations, an unwary receiver may still benefit from selective disclosure (cf. Proposition 1), although not to the same extend as wary receivers do.[45] However, when there are multiple competing senders,

[43] I.e., such 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).

[44] Our analysis with unwary receivers could also be adapted to study settings where receivers have rational expectations but there is uncertainty about the selectivity of disclosure, e.g., because receivers do not perfectly know the state of a sender's targeting technology as in Grubb (2011). For instance, there might be uncertainty about a sender's costs of acquiring information, $c_m \in \{c^L_m, c^H_m\}$ with $p_L := \Pr(c_m = c^L_m)$. Then, faced with a low-cost sender, rational receivers indeed underestimate the selectivity of disclosure when only low-cost senders acquire information, and our analysis applies with $N_m(U_m)$ now denoting the distribution of the receiver's perceived utility in the (hypothetical) case where the sender is low-cost with probability p_L and high-cost with probability $1-p_L$, and $\hat{S}_m(U_m)$ the respective distribution when costs are indeed low $c_m = c^L_m$. The case of a high-cost sender, where receivers now overestimate the selectivity of disclosure, would then require a separate analysis. Still, in what follows, we stick to the interpretation of unwary receivers.

[45] Concretely, while Property 2 ensures that wary receivers always benefit from selective disclosure, this is not the case for unwary receivers who might actually be worse off under selective disclosure when there is a monopolistic sender.

matters are substantially different, as we show next.

This becomes most transparent when senders are symmetric, that is when $F_m(u_m^i) = F(u_m^i)$. Then, as all senders will follow the same selective disclosure strategy in equilibrium and the respective utilities are drawn from identical distributions, the receiver's perception of any sender's option will be equally inflated. This is the key difference to the previously considered case with a monopolistic sender, where the receiver compared an inflated perception to the fixed value of an outside option (R).[46] In fact, when there is symmetry across senders, the respective distortions exactly cancel out and selective disclosure affects an unwary receiver exactly in the same way as a wary receiver. We next derive this insight more formally.

For this recall first that in equilibrium each sender discloses $u_{d_m} = \max \langle u_m^1, u_m^2 \rangle$. As both the true expected valuation, $u_{d_m} + E[u_m^i | u_m^i \leq u_{d_m}]$, as well as the one perceived by the unwary receiver, $u_{d_m} + E[u_m^i]$, are strictly increasing in u_{d_m}, and, by symmetry, for a given u_{d_m} constant over m, an unwary receiver's decision rule is the same as that of a wary receiver, namely to choose the option for which the disclosed value u_{d_m} is maximal. Note that this observation clearly makes use of symmetry across senders, as we further discuss below. We thus have the following striking implication:

Proposition 4 *Consider competition between at least two symmetric senders with unwary receivers. Then, all senders acquire information and disclose selectively, and the resulting outcome is the same, in terms of receiver decisions and receiver welfare, as if receivers were wary.*

Proposition 4 is, at the same time, both stronger and weaker than the corresponding results we obtained with wary receivers. It is weaker as we only argue that receivers surely benefit from a given sender's selective disclosure when also all other senders disclose selectively, thereby preserving symmetry.[47] But it is stronger because, when receivers are unwary, senders always disclose selectively, while this was not necessarily the case when wary receivers observed senders' strategies. In this sense, once we take into account senders' equilibrium strategies, unwary receivers can be strictly better off than wary receivers. Hence, receivers' ignorance is bliss—receivers reap the benefits of more informed decision making even though they make the wrong inferences. This immediately leads on to the following observations regarding policy.

When senders compete on a level playing field, Proposition 4 shows that such competition already sufficiently protects unwary receivers, providing them with the same advantage

[46] This case, in fact, corresponds to an extreme form of asymmetry across senders.

[47] When only some but not all senders disclose selectively, the insights from the analysis with a monopolistic sender still apply qualitatively, given that the value of some options is then inflated (more) than those of others, so that we have to trade off two types of errors (i.e., false positives and negatives).

of selective disclosure as wary receivers. Policy intervention that restricts selective disclosure would then be unwarranted, as in the case of wary receivers. What is now different however is that the previously discussed light-touch regulation of requiring consumer consent (to collect and use personalize data) would now no longer risk backfiring, as senders no longer benefit from committing not to selectively disclose information.

We conclude this section with some observations that apply when the economy contains a mix of wary and naive receivers. When senders can choose their information acquisition strategies individually for each receiver, then the preceding results apply (on a case-by-case logic) irrespective of whether all receivers are wary or unwary or whether there is a mixed composition of them. Suppose now instead that, while firms can disclose selectively to each receiver, their (now observed) investment in the collection of personalized data applies across all receivers. In this situation wary receivers can benefit from the presence of unwary receivers in case the (sufficiently large) presence of the latter induces senders to (observably) collect such data, while (notably for low M) they would not do so when the fraction of unwary receivers is smaller. Interestingly, this positive externality of unwary on wary consumers arises without hurting the former.[48]

Political Campaigning. While we so far focused on situations in which receivers take individual decisions, our equilibrium analysis and regulatory implications extend in a straightforward way also to cases of collective decision-making. To see this suppose, within an application to political campaigning, that $M = 2$ ex-ante symmetric candidates compete for voters. Denote by $U_m(z)$ voter z's expected utility when candidate m wins. In terms of motivation, 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. This generates scope for tailoring campaign messages to the political preferences of individual voters via selective disclosure.[49]

Clearly, this setting is identical to our previous analysis when the number of voters Z is equal to one. However, also for $Z > 1$, our previous analysis can be applied by noting that whenever a given voter is pivotal, his conditional expected utility equals $E[U^{(1)}]$, while it equals $E[U]$ else, as his vote then does not influence the collective decision. Hence, a

[48] See Armstrong (2015) for a survey of how the presence of both savvy (well informed and strategically sophisticated) as well as non-savvy consumers in a market can give rise to search externalities, according to which the non-savvy are protected by the savvy consumers, or ripoff externalities, which are present whenever the savvy consumers benefit from the presence of non-savvy types. Our externality is clearly different, as the presence of unwary consumers induces firms to choose (disclosure) strategies that also benefit wary consumers.

[49] This 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. See Alonso and Câmara (2016) for a model of optimal (rather than equilibrium) persuasion of voters in a setting with vertical (rather than horizontal) differentiation.

voter's ex-ante expected utility becomes

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

where $y=\binom{Z-1}{\frac{Z-1}{2}}\frac{1}{2}^{Z-1}$ denotes the probability with which any given voter will become pivotal.[50] Now, we already know that $E[U^{(1)}]$, and, thus, the term in braces is strictly higher when one of the candidates discloses selectively, and even more so when both candidates do. Through selective disclosure, individual votes better reflect the preferences of individual voters and it is in this sense that our previous analysis can be applied.[51]

Corollary 1 *Suppose Z voters decide by majority rule over $M=2$ ex-ante symmetric candidates. A switch from non-selective to selective disclosure by either (both) candidate(s) benefits each wary (unwary) voter, strictly so conditional on being pivotal. When candidates are unable to commit to their information acquisition strategy, in the unique equilibrium both choose to acquire information and then disclose selectively, maximizing voter utility. Privacy regulation restricting information acquisition by campaigners strictly reduces welfare.*

4 Equilibrium with Personalized Pricing

Within our main application to marketing, 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.[52] 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 less of a concern. Accordingly, this

[50] Note that, in equilibrium, the ex-ante likelihood of a vote for any of the two candidates must be equal to 1/2.

[51] While we cast our analysis into a framework where disclosure can be personalized for each individual receiver, the results also apply when communication is more coarse as it can be targeted (only) to groups of voters through the particular media that they consume. In this sense, our analysis also apply when candidates strategically adjust their messages to each individual channel in an increasingly fragmented media landscape.

[52] 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.

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 use this detailed information to charge personalized prices to customers.[53]

4.1 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 $U^{(1)} = \max_{m \in M} U_m$, and that, second, the price that the consumer pays is equal to the incremental utility relative to the second-highest such value, denoted by $U^{(2)}$. This is, thus, in the spirit of so-called mill pricing in a Hotelling model, where firms know the location of each consumer, which perfectly reflects the valuation for each individual product, and can make prices specific for each consumer location. Consequently, a consumer realizes the second order-statistic $U^{(2)}$.[54] 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 Proposition 3).

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

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

As the term in rectangular brackets is a convex function of U_m,[55] 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, any given firm m prefers to disclose selectively, irrespective of whether this is anticipated by the consumer or not. Thus, with personalized*

[53] The industrial organization literature on behavior-based price discrimination has focused on personalized pricing where, in particular, the past purchasing history of consumers is used; see, for example, Villas-Boas (1999) and Acquisti and Varian (2005). As we abstract from this dynamic feature, our analysis will be quite different.

[54] With a monopolistic seller, consumers are, thus, always pushed down to their reservation utility R.

[55] Its derivative with respect to U_m is $G^{(1:M\backslash m)}(U_m)$ which is increasing in U_m.

pricing it is the unique equilibrium outcome that all firms disclose selectively ($s_m = y$ for all $m \in M$).

4.2 Consumer Preferences with Personalized Pricing

From the perspective of consumers, the effect of selective disclosure depends now crucially on the degree of competition. Before we derive these results, note the stark contrast to the baseline case without personalized pricing, where the degree of competition affects firms' but not consumers' preferred choice of disclosure policy. The opposite holds with personalized pricing: the degree of competition affects consumers' but not firms' preferences with regards to selective disclosure.

As a starting point, consider a duopoly with $M = 2$, where the differences between the two cases are particularly stark. While without personalized pricing a consumer realized the maximum of the two expected utilities $U^{(1)} = \max \langle U_1, U_2 \rangle$, with personalized pricing the consumer now realizes the second-highest value, which for $M = 2$ is the minimum $U^{(2)} = \min \langle U_1, U_2 \rangle$. The consumer is now strictly worse off when any of the presently considered two firms discloses selectively. Formally, this can be seen in complete analogy to the argument for why without personalized pricing the consumer was strictly better off.

With personalized prices the consumer's expected utility is

$$E[U^{(2)}] = \int \left[\int \min \langle U_m, U_{m'} \rangle \, dG_{m'}(U_{m'}) \right] dG_m(U_m).$$

Given that the expression in rectangular brackets is a strictly concave function of U_m,[56] while it was a strictly convex function when without personalized pricing we applied the maximum, it is lower after a mean-preserving spread in $G_m(U_m)$. Intuitively, when firm m discloses selectively, a consumer's updating makes firms more differentiated from an ex-ante perspective. This ensures that in expectation the (winning) firm with the highest perceived value can extract a higher price. As we show next, however, this detriment to consumers from increased differentiation is reduced as M increases, when it becomes increasingly likely that each firm has a close competitor, in which case the efficiency benefits resulting from selective disclosure again accrue to the consumer, as stated in Lemma 2.

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 selectively, so that $S_m(U_m)$ instead of $N_m(U_m)$ applies. More precisely, consumers benefit from selective disclosure by firm m also under personalized pricing if and only if

$$\int [N_m(U_m) - S_m(U_m)] \left[G^{(2:M \setminus m)}(U_m) - G^{(1:M \setminus m)}(U_m) \right] dU_m > 0. \qquad (10)$$

[56] It can be written as $U_m - \int_{\underline{U}}^{U_m} [U_m - U_{m'}] dG_{m'}(U_{m'})$, with first derivative $1 - G_{m'}(U_m)$, which is decreasing.

Expression (10) 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 $G^{(2:M\setminus m)}(U_m) - G^{(1:M\setminus m)}(U_m)$. Furthermore, given single-crossing of $N_m(U_m)$ and $S_m(U_m)$, we can sign expression (10) unambiguously to be positive whenever there are sufficiently many firms (high M), as then, somewhat loosely speaking, both the first and the second highest value of all other $M-1$ utilities take on high realizations (that is, to the right-hand side of the rotation point). To express our results succinctly, we again impose symmetry across senders.[57]

Proposition 6 *Suppose that senders are symmetric. Then, with personalized pricing, wary consumers are indifferent between selective and non-selective disclosure by a monopolist ($M=1$). With a duopoly ($M=2$), consumers are always strictly worse off when a firm m switches to selective disclosure, regardless of the disclosure strategy of the rival firm. However, irrespective of the other firms' choices, consumers strictly benefit when any firm m chooses selective disclosure provided that there is sufficient competition (M large).*

Proposition 6 relates 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.

4.3 Implications for Regulation

We summarize the implications of our model for the regulation of firms' collection of personalized consumer data. When prices are personalized even wary consumers can be worse off under selective disclosure—in contrast to what we found when prices are not personalized. A second difference arises from firms' preferences. Given that with personalized pricing firms will always want to collect personal data and disclose selectively, regulation that requires consumers' consent can no longer backfire by granting firms commitment power, as in the baseline case. Regulation that strictly prohibits the collection and use of personal data, however, reduces efficiency and consumer welfare when competition is sufficiently intense.

[57] The subsequent results easily generalize to the case of asymmetric senders as long as distributions N_m and S_m have the same support across senders and sender types, as characterized by these distributions, are drawn from a finite set.

If consumers remain unwary of firms' capability to collect and use personally identifiable data for selective disclosure, consumers' perceived value for products can be inflated. While for the baseline analysis we showed that consumers may still be better off even with a monopolistic firm, with personalized pricing this is clearly no longer the case. Personalized pricing allows the firm to extract a consumer's *perceived* incremental utility relative to the next best choice available. If the perceived utility is inflated, this generates the potential for consumer exploitation. Facing a monopolistic firm, an unwary consumer would be better off staying out of the market.

Competition between senders, however, protects naive consumers, also with personalized pricing. As in the baseline case, this is most immediate when, under symmetry, the perception of the different offers is equally inflated, so that the decision of an unwary consumer fully matches that of a wary consumer. Interestingly, in this case an unwary consumer is now strictly better off than a wary consumer when personalized pricing is feasible, given that the price is reduced. This striking result follows because unwary consumers do not adjust expectations for firms' selection bias, which works towards reducing the perceived difference between the first-best and second-best alternative. Precisely, if $u^{(1)}$ is the highest disclosed fit (attribute) and $u^{(2)}$ is the second highest, an unwary consumer pays the price $\widehat{p} = u^{(1)} - u^{(2)}$, given that the expectations about the non-disclosed attribute of either firm wrongly remain unchanged at $E[u^i]$. 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. In other words, in the eyes of an unwary consumer, the firms offering the first-best and the second-best fit appear to be less differentiated, compared to the perceptions of wary consumers. We have thus the peculiar situation where any regulation that would *not* affect firms' strategies to gather and disclose information but would *only* increase consumers' awareness would not be in the interest of consumers, but in the interest of firms. Abstracting from these finer results, we summarize our implications for regulation as follows, based notably on Proposition 6:

Proposition 7 *When symmetric firms are able to not only selectively disclose information but also to price discriminate based on the personal data they collect, regulation that prohibits the data collection benefits consumers if and only if there is insufficient competition.*

5 Conclusion

The greater availability of personally identifiable data on the internet opens up new opportunities for tailoring advertising messages to the perceived preferences of particular consumers. Our analysis has broader implications also for face-to-face interactions between salespeople and consumers. 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 to use the limited time available (or the consumer's limited attention) to communicate only those product attributes that dovetail nicely with those preferences. Future research might develop a theory of the skills of good salespeople based on their ability to learn about consumers' preferences and to build up their sales talk accordingly. Furthermore, also old media may allow at least for a segmentation of receivers into coarse groups, so that different messages can be sent to groups with different preferences. Data collection on the internet naturally increases firms' ability to target their communication to individual receivers.

To summarize, in our baseline setting with individual decisions at fixed prices, wary receivers benefit from selective disclosure for a broad set of distributions satisfying log-concavity. However, senders' incentives to become better informed—as a basis of (more) selective disclosure—are 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. Even though senders are forced by their own incentives to become better informed, (more) selective disclosure ends up either benefitting or hurting them in equilibrium. Notably, senders benefit from increased information only when competition is intense. Policy intervention that makes information acquisition *observable*, for example by requiring consumer consent, may end up hurting consumers because by not requiring consent senders are able to commit not to acquire information.

The introduction of personalized pricing changes the outcome of our baseline analysis in one important way.[58] The extent to which the efficiency gains associated with more informative communication are shared between firms and consumers depends now on the degree of competition. As selective disclosure based on better information is also found to dampen competition by increasing perceived differentiation, there is a trade-off from the perspective of consumers. Policy intervention is notably not warranted when there is sufficient competition. For policy purposes we also consider the case where (some) receivers remain naive about senders' capabilities, thereby not properly discounting their valuation

[58] Such price discrimination may only be feasible for services or low-value products, when customers or intermediaries have little scope for arbitrage.

for the adverse selection that is implicit in the fact that the disclosed attribute is the most favorable. For a particularly clear-cut case, with personalized pricing and insufficient competition naive consumers risk being exploited and would be better off staying out of the market. On the other hand, sufficient competition allows naive receivers to benefit from selective disclosure as much, and sometimes even more, as wary consumers. The key to the last observation is that when senders are equally positioned to selectively disclose, the effect of inflated values effectively cancels out.

We obtained our results in a stylized model where senders can freely acquire and divulge information only about their own offerings in a private-value environment.[59] An extension could allow senders to disclose information about their competitors. Information could be costly, as in the law and economics literature on transparency. Information could be also sold by an information broker as in Taylor (2004), Bergemann and Bonatti (2015), and Montes, Sand-Zantman, and Valletti (2019). Finally, our baseline comparison of the impact of selective disclosure on receiver and sender preferences is relevant also beyond our main application to selective disclosure based on personalized data, e.g., when considering the selective disclosure of vertical product dimensions such as quality test results. Exploration of these avenues is left for future work.

6 References

Acquisti A, Varian HR (2005) Conditioning prices on purchase history. *Marketing Sci.* 24(3):367–381.

Agrawal A, Gans J, Goldfarb A (2018) *Prediction Machines* (Harvard Business Review Press, Boston, MA).

Alonso R, Câmara O (2016) Persuading voters. *Am. Econ. Rev.* 106(11):3590–3605.

American Association of Advertising Agencies (2009) *Self-Regulatory Principles for Online Behavioral Advertising.* Accessed July 13, 2019, https://www.iab.com/wp-content/uploads/2015/05/ven-principles-07-01-09.pdf.

An M (1998) Logconcavity versus logconvexity: A complete characterization. *J. Econ. Theory* 80(2):350–369.

Anderson S, de Palma A (2012) Competition for attention in the information (overload) age. *RAND J. Econ.* 43(1):1–25.

[59] 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.

Armstrong M (2015) Search and ripoff externalities. *Rev. Ind. Organ.* 47(3):273–302.

Athey S, Gans J (2010) The impact of targeting technology on advertising and media competition. *Am. Econ. Rev.* 100(2):608–613.

Bagnoli M, Bergstrom T (2005) Log-concave probability and its applications. *Econ. Theory* 26(2):445–469.

Baricz Á (2010) Geometrically concave univariate distributions. *J. Math. Anal. Appl.* 363(1):182–196.

Bergemann D, Bonatti A (2015) Selling cookies. *Am. Econ. J. Microecon.* 7(3):259–294.

Bhattacharya S, Mukerjee A (2013) Strategic information revelation when experts compete to influence. *RAND J. Econ.* 44(3):522–544.

Board S (2009) Revealing information in auctions: The allocation effect. *Econ. Theory* 38(1):125–135.

Board S, Lu J (2018) Competitive information disclosure in search markets. *J. Polit. Econ.* 126(5):1965–2010.

Butters GR (1977) Equilibrium distributions of sales and advertising prices. *Rev. Econ. Stud.* 44(3):465–491.

Caillaud B, Tirole J (2007) Consensus building: How to persuade a group. *Am. Econ. Rev.* 97(5):1877–1900.

Campbell J, Goldfarb A, Tucker C (2015) Privacy regulation and market structure. *J. Econ. Manag. Strategy* 24(1):47–73.

Casadesus-Masanell R, Hervas-Drane A (2015) Competing with privacy. *Manag. Sci.* 61(1):229–246.

Chakraborty A, Harbaugh R (2007) Comparative cheap talk. *J. Econ. Theory* 132(1):70–94.

Chakraborty A, Harbaugh R (2010) Persuasion by cheap talk. *Am. Econ. Rev.* 100(5):2361–2382.

Che Y-K, Dessein W, Kartik N (2013) Pandering to persuade. *Am. Econ. Rev.* 103(1):47–79.

DellaVigna S, Gentzkow M (2010) Persuasion: Empirical evidence. *Annu. Rev. Econ.* 2:643–669.

Dessein W, Galeotti A, Santos T (2016) Rational inattention and organizational focus. *Am. Econ. Rev.* 106(6):1522–1536.

Eyster E, Rabin M (2005) Cursed equilibrium. *Econometrica*, 73(5):1623–1672.

Farrell J (1986) Voluntary disclosure: Robustness of the unraveling results and comments on its importance. Greison RE, ed. *Antitrust and Regulation* (Lexington Books, Lexington, MA).

Federal Trade Commission (2012) *Protecting Consumer Privacy in an Era of Rapid Change. Recommendations for Businesses and Policy Makers.* Accessed July 13, 2019, http://www.ftc.gov/os/2012/03/120326privacyreport.pdf.

Fishman MJ, Hagerty KM (1990) The optimal amount of discretion to allow in disclosure. *Q. J. Econ.* 105(2):427–444.

Ganuza J-J, Penalva JS (2010) Signal orderings based on dispersion and the supply of private information in auctions. *Econometrica* 78(3):1007-1030.

Gentzkow M, Kamenica E (2017) Competition in persuasion. *Rev. Econ. Stud.* 84(1):300–322.

Glazer J, Rubinstein A (2004) On optimal rules of persuasion. *Econometrica* 72(6):1715–1736.

Goldfarb A, Tucker C (2011) Privacy regulation and online advertising. *Manag. Sci.* 57(1):57–71.

Grossman S (1981) The informational role of warranties and private disclosure about product quality. *J. Law Econ.* 24(3):461–483.

Grubb MD (2011) Developing a reputation for reticence. *J. Econ. Manag. Strategy* 20(1):225–268.

Hermalin B, Katz M (2006) Privacy, property rights, and efficiency: The economics of privacy as secrecy. *Quant. Mark. Econ.* 4(3):209–239.

Hirshleifer DA, Lim SS, Teoh SH (2004) Disclosure to an audience with limited attention. SSRN Working Paper.

Hirshleifer J (1971) The private and social value of information and the reward to inventive activity. *Am. Econ. Rev.* 61(4):561–574.

Holmström B (1999) Managerial incentive problems: A dynamic perspective. *Rev. Econ. Stud.* 66(1):169–182.

Johnson JP (2013) Targeted advertising and advertising avoidance. *RAND J. Econ.* 44(1):128–144.

Johnson JP, Myatt DP (2006) On the simple economics of advertising, marketing, and product design. *Am. Econ. Rev.* 96(3):756–784.

Kamenica E, Gentzkow M (2011) Bayesian persuasion. *Am. Econ. Rev.* 101(6):2590–2615.

Kamenica E, Mullainathan A, Thaler R (2011) Helping consumers know themselves. *Am. Econ. Rev. Papers Proc.* 101(3):417–422.

Lancaster K (1966) A new approach to consumer theory. *J. Polit. Econ.* 74(2):132–157.

Lewis TR, Sappington DEM (1994) Supplying information to facilitate price discrimination. *Intern. Econ. Rev.* 35(2):309–327.

Milgrom P (1981) Good news and bad news: Representation theorems and applications. *Bell J. Econ.* 12(2):380–391.

Milgrom P, Roberts J (1986) Relying on information of interested parties. *RAND J. Econ.* 17(1):18–32.

Montes R, Sand-Zantman W, Valletti T (2019) The value of personal information in online markets with endogenous privacy. *Manag. Sci.* 65(3):955–1453.

Posner RA (1981) The economics of privacy. *Am. Econ. Rev.* 71(2):405–409.

Prékopa A (1973) On logarithmic concave measures and functions. *Acta Sci. Math.* 34:335–343.

Rayo L, Segal I (2010) Optimal information disclosure. *J. Polit. Econ.* 118(5):949–987.

Shaked M, Shanthikumar G (2007) *Stochastic Orders* (Springer, New York, NY).

Shapiro C, Varian HR (1997) US government information policy. Working paper, University of California, Berkeley.

Shen Q, Villas-Boas JM (2018) Behavior-based advertising. *Manag. Sci.* 64(5):2047–2064.

Shavell S (1994) Acquisition and disclosure of information prior to sale. *RAND J. Econ.* 25(1):20–36.

Simon HA (1971) Designing organizations for an information-rich world. Greenberger M, ed. *Computers, Communication, and the Public Interest* (The Johns Hopkins Press, Baltimore, MD), 38–72.

Sims CA (2003) Implications of rational inattention. *J. Monetary Econ.* 50(3):665–690.

Stigler GJ (1980) An introduction to privacy in economics and politics. *J. Leg. Stud.* 9(4):623–644.

Taylor C (2004) Consumer privacy and the market for customer information. *RAND J. Econ.* 35(4):631–651.

Tucker CE (2012) The economics of advertising and privacy. *Int. J. Ind. Organ.* 30(3):326–329.

Van Zandt T (2004) Information overload in a network of targeted communication. *RAND J. Econ.* 35(3):542–560.

Villas-Boas MJ (1999) Dynamic competition with customer recognition. *RAND J. Econ.* 30(4):604–631.

White House (2012) *Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy.* Accessed July 13, 2019, https://www.hsdl.org/?abstract&did=700959.

Whitt W (1985) Uniform conditional variability ordering of probability distributions. *J. Appl. Probab.* 22(3):619–633.

Wu S, Debnath L (2009) A generalization of l'Hôspital-type rules for monotonicity and its application. *Appl. Math. Lett.* 22(2):284–290.

Zeckhauser RJ, Marks DVP (1996) Sign posting: The selective revelation of product information. Zeckhauser RJ, Keeney RL, Sebenius JK, eds. *Wise Choices: Games, Decisions, and Negotiations* (Harvard Business School Press, Boston, MA), 22–41.

Appendix A Proofs

A.1 Impact of Selective Disclosure

In this Appendix we collect the results and proofs omitted from Section 2; for additional material and verification of the examples see Appendix B.2.

Proof of Proposition 1. Suppose that $R \in [2E[u], \overline{u} + E[u]]$. Slice the shaded regions A, B, Y, and Z into iso-payoffs symmetrically with respect to R, as shown in Figure 4. For every offset level $\delta > 0$ draw two iso-payoffs on either side of R, with equations $u^2 = R - \delta - u^1$ and $u^2 = R + \delta - u^1$.

First, consider the iso-payoffs for which $\delta \in [0, R - (\overline{u} + \underline{u})]$ in areas B and Y. To show that the expected gain in Y is higher than the expected loss in B, on the iso-payoff of level $R - \delta$ take a generic point $b \in B$ with coordinates $(u^1 = -\delta - \varepsilon + E[u], u^2 = R - E[u] + \varepsilon)$ where $\varepsilon \in [0, \overline{u} - R + E[u]]$ is the vertical distance of u^2 from $R - E[u]$. The expected loss relative to R at this point is

$$-\delta f(-\delta - \varepsilon + E[u]) f(R - E[u] + \varepsilon). \tag{11}$$

At the corresponding point $y \in Y$, with coordinates $(u^1 = \delta - \varepsilon + E[u], u^2 = R - E[u] + \varepsilon)$ on the iso-payoff of level $R + \delta$, the expected gain relative to R is

$$\delta f(\delta - \varepsilon + E[u]) f(R - E[u] + \varepsilon). \tag{12}$$

Note that the horizontal distance of y to $E[u]$ is lower than the horizontal distance of b to $E[u]$, i.e., $|(\delta - \varepsilon + E[u]) - E[u]| \leq |E[u] - (-\delta - \varepsilon + E[u])|$ given that $\varepsilon \geq 0$. If F is unimodal, the density decreases the farther away the realization is from $E[u]$, thus, using symmetry we have $f(-\delta - \varepsilon + E[u]) \leq f(\delta - \varepsilon + E[u])$. Summing the gains (12) and losses (11) over δ and ε in regions Y and B, we conclude that the expected gain is greater than the expected loss

$$\int_0^{R-(\overline{u}+\underline{u})} \int_0^{\overline{u}-R+E[u]} \delta f(R - E[u] + \varepsilon) [f(E[u] - \varepsilon + \delta) - f(E[u] - \varepsilon - \delta)] \, d\varepsilon d\delta \geq 0. \tag{13}$$

Second, construct paired iso-payoffs with offset $\delta \in [R - (\overline{u} + \underline{u}), E[u] - \underline{u}]$ in regions A and Z. The following one-to-one function maps each point in A to each point in Z

$$(E[u] - \delta - \varepsilon, R - E[u] + \varepsilon) \to (R - E[u] - \varepsilon, E[u] + \delta + \varepsilon),$$

where $\varepsilon \in [0, E[u] - \underline{u} - \delta]$, illustrated by the stars in Figure 4. To compare the density of points in A and Z, we map all points (such as the original star) of A to points in X (such as the dot within X marked in Figure 4) through the function $(E[u] - \delta - \varepsilon, R - E[u] + \varepsilon) \to (E[u] + \delta + \varepsilon, R - E[u] + \varepsilon)$. Given that region X is symmetric to region A with respect to $E[u]$ and that F is symmetric, this first mapping preserves the density of the

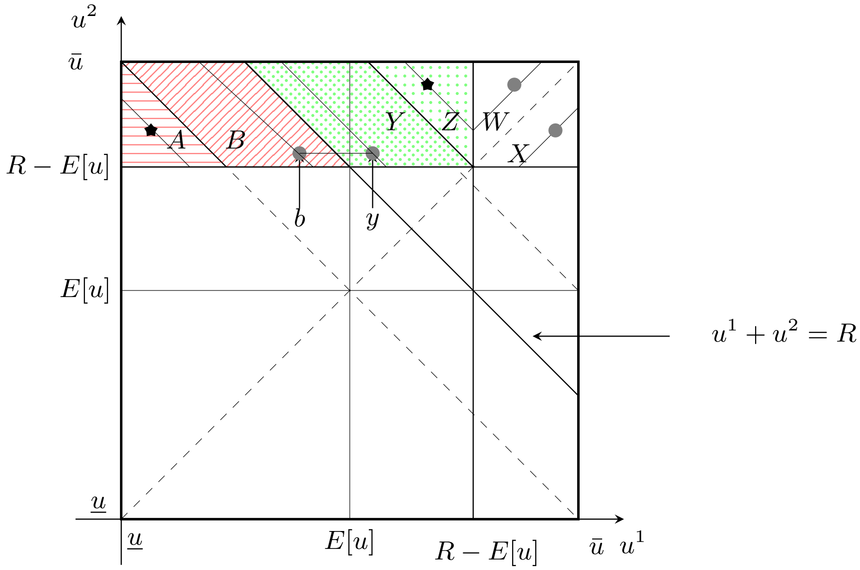

Figure 4: Welfare impact on unwary receiver.

points. Second, map each point in X to a point in W through the inverse function $(E[u] + \delta + \varepsilon, R - E[u] + \varepsilon) \rightarrow (R - E[u] + \varepsilon, E[u] + \delta + \varepsilon)$. Given that u^1 and u^2 are i.i.d., the density of the points is preserved. Thus, the initial star in A has the same density as the final dot in W. Comparing the density of points in A and Z is therefore equivalent to comparing the density of points in W and Z. The density of the star in Z is higher than the density of the dot in W because it is closer to $E[u]$: $f(R - E[u] - \varepsilon) \geq f(R - E[u] + \varepsilon)$. We conclude that the expected gain is greater than the expected loss

$$\int_{R-(\overline{u}+\underline{u})}^{E[u]-\underline{u}} \int_0^{E[u]-\underline{u}-\delta} \delta f(\delta + E[u] + \varepsilon)\left[f(R - E[u] - \varepsilon) - f(R - E[u] + \varepsilon)\right] d\varepsilon d\delta \geq 0. \tag{14}$$

We have established that if $R \in [2E[u], \overline{u} + E[u]]$ and the distribution F is unimodal, the unwary receiver is better off under selective than non-selective disclosure.[60]

To complete the proof, note that if $R \geq \overline{u} + E[u]$ the receiver never accepts the

[60] Alternatively, to compare the density of points in regions A and Z we calculate the radiuses of the circles, with center $(E[u], E[u])$, that pass through two corresponding points. The i.i.d. and symmetry assumptions imply that points on the same circle have the same density. Unimodality then implies that points with a higher distance from the center of Figure 4 have a lower density. The radius of a point in A is

$$(E[u] - \delta - \varepsilon - E[u])^2 + (R - E[u] + \varepsilon - E[u])^2,$$

while the radius of a point in Z is

$$(R - E[u] - \varepsilon - E[u])^2 + (E[u] + \delta + \varepsilon - E[u])^2.$$

Thus, the density of a point in Z is higher than the corresponding point in A if and only if $4E[u]\varepsilon - 2\varepsilon R \leq 2\varepsilon R - 4E[u]\varepsilon$, i.e., $R \geq 2E[u]$.

offering and thus always obtains utility R regardless of the disclosure regime. If, instead, $R \leq 2E[u]$, the points corresponding to losses (to the left of the iso-payoff of level R) are now closer to $E[u]$ and thus become more likely than those corresponding to gains (with the same offset level to the right of the iso-payoff). Because of this reversal in the distance of the points from $E[u]$, if F is unimodal and $R \leq 2E[u]$, the unwary receiver is better off under non-selective disclosure than under selective disclosure. **Q.E.D.**

A.2 Equilibrium Analysis

In this Appendix we collect all omitted proofs from Sections 3 and 4.

Proof of Lemma 2. The receiver's expected utility can be written as

$$E[U^{(1)}] = \int \left[\int \max \left\langle U^{(1:M\backslash m)}, U_m \right\rangle dG^{(1:M\backslash m)}(U^{(1:M\backslash m)}) \right] dG_m(U_m). \tag{15}$$

Given that the expression in brackets is a convex function of U_m,[61] it is higher after a mean-preserving spread in $G_m(U_m)$. **Q.E.D.**

Proof of Lemma 3. For any given distribution of the receiver's next best alternative $G^{(1:M\backslash m)}(\cdot)$, the difference in the likelihood that option m is chosen when sender m switches from $N_m(\cdot)$ to $S_m(\cdot)$ is given by[62]

$$\Delta q_m = \int_{\underline{U}}^{\overline{U}} G^{(1:M\backslash m)}(U_m) d\left[S_m(U_m) - N_m(U_m)\right] = \int_{\underline{U}}^{\overline{U}} Z_m(U_m) dG^{(1:M\backslash m)}(U_m), \tag{16}$$

where, with $Z_m(U_m) = N_m(U_m) - I_m(U_m)$, the second line follows from integration by parts. Now consider two choices for the distribution $G^{(1:M\backslash m)}(\cdot)$: $H'(\cdot)$ and $H''(\cdot)$, where the latter dominates in the likelihood ratio order. Further, denote the support of H' by $\left[\underline{U}', \overline{U}'\right]$ and by $\left[\underline{U}'', \overline{U}''\right]$ the respective support of H'', where from the likelihood ratio property we must have that $\underline{U}' \leq \underline{U}''$ as well as $\overline{U}' \leq \overline{U}''$. We now apply the following

[61] It can be written as $U_m + \int_{U_m}^{\infty} [U^{(1:M\backslash m)} - U_m] dG^{(1:M\backslash m)}(U^{(1:M\backslash m)})$ with first derivative $G^{(1:M\backslash m)}(U_m)$, which is increasing.

[62] The upper and lower bounds of the integral are chosen to contain the supports of $N_m(\cdot)$ and $S_m(\cdot)$.

transformations, focusing on the non-trivial case where $\widetilde{U}_m$ satisfies $\underline{U}'' \leq \widetilde{U}_m \leq \overline{U}'$:[63]

$$\begin{aligned}
\Delta q''_m &= \int_{\underline{U}'}^{\overline{U}''} Z_m(U_m)h''(U_m)dU_m \\
&= \int_{\underline{U}'}^{\widetilde{U}_m} Z_m(U_m)\frac{h''(U_m)}{h'(U_m)}h'(U_m)dU_m + \int_{\widetilde{U}_m}^{\overline{U}''} Z_m(U_m)\frac{h''(U_m)}{h'(U_m)}h'(U_m)dU_m \\
&> \frac{h''(\widetilde{U}_m)}{h'(\widetilde{U}_m)}\int_{\underline{U}'}^{\overline{U}'} Z_m(U_m)h''(U_m)dU_m = \frac{h''(\widetilde{U}_m)}{h'(\widetilde{U}_m)}\Delta q'_m,
\end{aligned}$$

where, for the second line, we have used that $Z_m(U_m) < 0$ for $U_m < \widetilde{U}_m$ and $Z_m(U_m)$ for $U_m > \widetilde{U}_m$ and, for the third line, that $\frac{h''(U_m)}{h'(U_m)}$ is increasing in U_m. **Q.E.D.**

Proof of Proposition 3. Given Lemma 3 it only remains to prove that such a (finite) cutoff M' indeed exists. Take some $U \in (\widetilde{U}, \overline{U})$ (where we have dropped the subscript due to sender symmetry). Then for arbitrary small, but strictly positive, $\varepsilon > 0$ there exists a finite boundary $\widehat{M}$ such that for all $M > \widehat{M}$ we have, by construction, that $G^{(1:M\setminus m)}(U) = S^{M_S}(U)N^{M-M_S-1}(U) < \varepsilon$, where M_S denotes the number of firms other than m that disclose selectively. I.e., almost all mass for the second-best alternative lies above the considered value $U > \widetilde{U}$. The assertion follows then immediately from expression (16) in the preceding proof and the single-crossing mean-preserving spread between $N(U)$ and $S(U)$. **Q.E.D.**

Proof of Proposition 6. It remains to show the result for large M. To simplify the exposition, without loss of generality we consider the choice of firm $m = 1$ and thus a switch from $G_1(U_1) = N_1(U_1)$ to $G_1(U_1) = S_1(U_1)$. 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

$$\int_{\underline{U}}^{\overline{U}} v_1(U_1)d\left[S_1(U_1) - N_1(U_1)\right]. \tag{17}$$

Next, note that $v_1(U_1)$ has derivative

$$\eta(U_1) = G^{(2:M\setminus 1)}(U_1) - G^{(1:M\setminus 1)}(U_1) = \sum_{m\in M\setminus 1}\left([1 - G_m(U_1)]\prod_{m'\notin\{1,m\}} G_{m'}(U_1)\right).$$

[63] Note that from the restrictions that condition (8) imposes on $H'(\cdot)$ and $H''(\cdot)$ it holds that $N_m(\cdot)$ and $S_m(\cdot)$ have strictly positive mass for $U_m \in \left[\underline{U}'', \overline{U}'\right]$. Then, if $\underline{U}'' > \widetilde{U}_m$, it is immediate that $S_m(\cdot)$ first-order stochastically dominates $N_m(\cdot)$ on the relevant support, such that we must have $\Delta q''_m > 0$; similarly, if $\overline{U}' < \widetilde{U}_m$, we must have that $\Delta q'_m < 0$.

So, using integration by parts, we can transform (17) to obtain

$$\int_{\underline{U}}^{\overline{U}} \eta(U_1)[N_1(U_1) - S_1(U_1)]dU_1. \tag{18}$$

We now drop the respective subscripts due to sender symmetry and define $\phi(U) = \eta(U)/\eta(\widetilde{U})$. (Recall that $\widetilde{U}$ denotes the rotation point.) Then, extending the expression in (18), for each realization of U, by multiplying and dividing with the term $\eta(\widetilde{U})$, and noting that $\eta(\widetilde{U}) > 0$, a sufficient condition for (18) to be greater than zero is that

$$\int_{\underline{U}}^{\widetilde{U}} [N(U) - S(U)]\, \phi(U)dU + \int_{\widetilde{U}}^{\overline{U}} [N(U) - S(U)]\, \phi(U)dU > 0. \tag{19}$$

Then, noting that for $U < \widetilde{U}$ it holds that $\phi(U) \leq \frac{1-N(U)}{1-N(\widetilde{U})}\left(S(U)/S(\widetilde{U})\right)^{M-2}$, the first (negative) integral in (19) converges to zero as $M \to \infty$, regardless of whether $G_m = N$ or $G_m = S$ for any other sender $m \neq 1$. It remains to show that the second (positive) integral remains bounded away from zero, which follows immediately as, for any $\widetilde{U} < U < \overline{U}$, we can write $\phi(U) \geq \frac{1-N(U)}{1-N(\widetilde{U})}\left(N(U)/N(\widetilde{U})\right)^{M-2} > 0$. **Q.E.D.**

Appendix B Supplementary Material

B.1 Alternative Foundation for Horizontal Differentiation

In parallel to the discussion in Section 3.1, we now briefly present an alternative foundation for receiver preferences in terms of the standard location approach to horizontal differentiation. For each component of the sender's offering, the receiver's willingness to pay is determined by how much the actual specification (hard facts) differs from the receiver's preferred specification (preferences). Concretely, suppose a sender's offering has two attributes, given by x^i, $i = 1, 2$, which are, from an ex-ante perspective, each independently and uniformly distributed on a Salop circle of circumference two. Each receiver has a preferred location for each attribute, y^i, $i \in \{1, 2\}$, which is also independently and uniformly distributed on the corresponding circle. The willingness to pay for each attribute is decreasing in the distance $\delta^i := |x^i - y^i|$ and given by $u^i = \alpha - \delta^i$ for some $\alpha > 1$, so that u^i is distributed uniformly over $[\alpha - 1, \alpha]$. The uniform distributions of x^i and y^i ensure that ex-ante, i.e., given knowledge of either x^i or y^i, the u^i are identically distributed so that differentiation is indeed only horizontal.

B.2 Impact of Selective Disclosure

Verification of Material Reported in Text. Here we collect additional material relevant for the analysis and examples reported in Sections 2.2 and 2.3.

Verification of Observation 1. For all R the difference between the receiver's values under selective and non-selective disclosure is

$$V_S - V_N = \frac{(V_S - V_{\hat{S}}) + (V_S - V_{\hat{T}})}{2} > \frac{(V_S - V_{\hat{T}})}{2} > 0, \tag{20}$$

where the first inequality follows from $V_S - V_{\hat{S}} > 0$ and the second inequality uses the assumption $V_S > V_{\hat{T}}$.

Verification of Observation 1*. Suppose by contradiction that at some R we have $V_{\hat{S}} > V_{\hat{T}}$. Then at that R, (20) holds, i.e., $V_S > V_N$. This contradiction proves the claim.

Example B.1 (Exponential) *The unwary receiver's expected payoff is always (for all R) higher under selective rather than non-selective disclosure, $V_{\hat{S}} > V_N$, if u is exponentially distributed, $f(u) = \lambda e^{-\lambda u}$ with $\lambda > 0$ and support $[0, \infty)$.*

Verification of Example B.1. Evaluating $V_N = \int_0^\infty \max \langle R, u + E[u] \rangle \, dF(u)$ and $V_{\hat{S}} = RF(R - E[u])^2 + \int_{R-E[u]}^\infty [u_d + E[u|u \le u_d]] \, dF(u_d)^2$ using $F(u) = 1 - e^{-\lambda u}$, $E[u] = 1/\lambda$ and $E[u|u \le u_d] = 1/\lambda - u_d e^{-\lambda u_d} / \left(1 - e^{-\lambda u_d}\right)$, we obtain after some transformations

$$V_N - V_{\hat{S}} = e^{-\lambda\left(R - \frac{1}{\lambda}\right)} \left[Re^{-\lambda\left(R - \frac{1}{\lambda}\right)} - \frac{1}{\lambda} \right].$$

The result then follows from noting that $Re^{-\lambda\left(R-\frac{1}{\lambda}\right)}$, with $R \geq 1/\lambda$, is maximized at $R = 1/\lambda$ where it takes on value $1/\lambda$.

We next establish the result that S is a clockwise rotation of N for a large set of commonly used distributions F that satisfy logconcavity. To reframe the problem, note that observation of $x = \max\langle u^1, u^2\rangle$ under selective disclosure corresponds to a crossing point between S and N if and only if $F(x)^2 = F(\mathcal{U}(x) - E[u])$.[64] Because F is strictly increasing and continuous, we can conveniently define the function

$$\gamma(x) := F^{-1}\left(\sqrt{F(\mathcal{U}(x) - E[u])}\right) \text{ for } x \in \left(\mathcal{U}^{-1}(\underline{u} + E[u]), \overline{u}\right),$$

whose fixed points exactly correspond to crossings between S and N. In the following Proposition we exploit this property of function γ to characterize the set of distributions F satisfying clockwise rotation.

Proposition B.1 *The distribution S of $E[u^1 + u^2 | \max\langle u^1, u^2\rangle]$, the expected sum given the maximum, crosses only once and from above the distribution N of the expected sum given a single variable, $E[u_1 + u_2 | u_1]$, in the interior of the support if and only if:*
(A) the left-hand integral $L(u) = \int_{\underline{u}}^{u} F(y)\,dy$ is logconcave at the upper bound and
(B) the function γ has a unique fixed point in the interior of the domain $\left(\mathcal{U}^{-1}(\underline{u} + E[u]), \overline{u}\right)$.

Proof of Proposition B.1. We first establish that the left-hand integral $L(u) = \int_{\underline{u}}^{u} F(y)\,dy$ is (strictly) logconcave at the upper bound $\Leftrightarrow x < \gamma(x)$, or equivalently $S(U) < N(U)$, in a left neighborhood of the upper bound of the support. To prove this, note that there exists a value x_l such that $x < \gamma(x)$ for all $x > x_l$ if and only if at the upper bound $\left.\frac{d\gamma(x)}{dx}\right|_{x=\overline{u}} < 1$. From $\gamma(x) = F^{-1}\left(\sqrt{F(\mathcal{U}(x) - E[u])}\right)$ and $\mathcal{U}(x) = 2x - \frac{L(x)}{F(x)}$ we have

$$\begin{aligned} \frac{d\gamma(x)}{dx} &= \frac{\mathcal{U}'(x)}{2} \cdot f(\mathcal{U}(x) - E[u]) \cdot \frac{1}{2}\left(F(\mathcal{U}(x) - E[u])\right)^{-\frac{1}{2}} \cdot \frac{1}{f(\gamma(x))} \\ &= \frac{\mathcal{U}'(x)}{2} \cdot \frac{\frac{f(\mathcal{U}(x)-E[u])}{F(\mathcal{U}(x)-E[u])}}{\frac{f(\gamma(x))}{F(\gamma(x))}} \end{aligned} \tag{21}$$

where for the second equality we have multiplied and divided by $F(\mathcal{U}(x) - E[u])$ and used the fact that $\left(F(\mathcal{U}(x) - E[u])\right)^{\frac{1}{2}} = F\left(F^{-1}\left(\sqrt{F(\mathcal{U}(x) - E[u])}\right)\right) = F(\gamma(x))$. Since at the upper bound $\gamma(\overline{u}) = \overline{u} = \mathcal{U}(\overline{u}) - E[u]$, we have

$$\frac{d\gamma(\overline{u})}{dx} = \frac{\mathcal{U}'(\overline{u})}{2} \cdot 1 = \frac{1 + \frac{L(\overline{u})f(\overline{u})}{F(\overline{u})^2}}{2}.$$

[64] To any crossing point $\tilde{U}$ between S and N, where by definition $N(\tilde{U}) = S(\tilde{U})$, there correspond two realizations $x = \max\langle u^1, u^2\rangle$ and u^i of the disclosed variable that induce receiver's expected utility $\tilde{U} = \mathcal{U}(x)$ and $\tilde{U} = u^i + E[u]$, respectively under selective and non-selective disclosure. Thus, x corresponds to a crossing point between N and S if and only if $S(\mathcal{U}(x)) = F(x)^2 = F(\mathcal{U}(x) - E[u]) = N(\mathcal{U}(x))$.

Thus $\frac{d\gamma(\overline{u})}{dx} \leq 1 \Leftrightarrow L$ is logconcave at the upper bound.

Second, by construction γ has a unique fixed point in the interior of the domain $\left(\mathcal{U}^{-1}\left(\underline{u}+E\left[u\right]\right),\overline{u}\right)$ if and only if there is a unique crossing between N and S in the interior of their common support, i.e. if and only if S is a rotation of N. By the observation in the first paragraph, the left-hand integral $L\left(u\right) = \int_{\underline{u}}^{u} F\left(y\right)dy$ is (strictly) logconcave at the upper bound if and only if $x < \gamma\left(x\right)$ for all x sufficiently high (i.e. at the last interior fixed point γ must cross the 45-degree line from below so that $S\left(U\right) < N\left(U\right)$ in a left neighborhood of the upper bound of the support). Thus, conditions (A) and (B) hold if and only if S is clockwise rotation of N. **Q.E.D.**

Logconcavity of the left-hand integral L is a relatively weak condition, which is implied by logconcavity of the distribution F, in turn implied by logconcavity of the density f.[65] As verified in below, a large class of distributions satisfying Proposition B.1's condition (A), logconcavity of L, also satisfy condition (B). In fact, we show next that under Proposition B.1's condition (A), logconcavity of L at the upper bound, concavity of γ implies condition (B), uniqueness of the interior fixed point of γ. Indeed, if γ is concave, it can cross the 45-degree line at most twice and, if there are exactly two crossings, the first crossing is from below and the second from above. Combined with the law of iterated expectations (implying existence of an interior crossing) and logconcavity of L at the upper bound (implying that at the last crossing γ crosses the 45-degree line from below), concavity of γ is sufficient for uniqueness of the interior fixed point of γ. Differentiating (21), γ is concave if

$$\frac{d^2\gamma\left(x\right)}{dx^2} = \underbrace{\frac{\mathcal{U}''\left(x\right)}{2}}_{\leq 0 \Leftrightarrow \frac{L(x)}{F(x)} \textit{ is convex}} \frac{\frac{f(\mathcal{U}(x)-E[u])}{F(\mathcal{U}(x)-E[u])}}{\frac{f(\gamma(x))}{F(\gamma(x))}} + \underbrace{\frac{\mathcal{U}'\left(x\right)}{2}}_{\geq 0} \underbrace{\frac{d}{dx}\left(\frac{\frac{f(\mathcal{U}(x)-E[u])}{F(\mathcal{U}(x)-E[u])}}{\frac{f(\gamma(x))}{F(\gamma(x))}}\right)}_{\leq 0 \text{ if } F \textit{ is logconcave \& } \frac{f'F}{f^2} \textit{ is increasing.}} \leq 0. \quad (22)$$

A first force leading toward concavity of γ is concavity of $\mathcal{U}$ defined in (6), equivalent to convexity of the mean-advantage-over-inferiors L/F. Given that $\mathcal{U}'\left(x\right) > 0$, the second addend is also negative whenever

$$\frac{f'(\mathcal{U}(x)-E[u])F(\mathcal{U}(x)-E[u])-f(\mathcal{U}(x)-E[u])^2}{F(\mathcal{U}(x)-E[u])^2}\mathcal{U}'\left(x\right)\frac{f(\gamma(x))}{F(\gamma(x))} \leq \frac{f(\mathcal{U}(x)-E[u])}{F(\mathcal{U}(x)-E[u])}\frac{f'(\gamma(x))F(\gamma(x))-f(\gamma(x))^2}{F(\gamma(x))^2}\gamma'\left(x\right). \quad (23)$$

Using (21) and logconcavity of F, (23) is equivalent to

$$\left(1-\frac{f'(\mathcal{U}(x)-E[u])F(\mathcal{U}(x)-E[u])}{f(\mathcal{U}(x)-E[u])^2}\right)+\left(\frac{f'(\gamma(x))F(\gamma(x))}{f(\gamma(x))^2}-\frac{f'(\mathcal{U}(x)-E[u])F(\mathcal{U}(x)-E[u])}{f(\mathcal{U}(x)-E[u])^2}\right) \geq 0.$$

Given that $\mathcal{U}\left(x\right)-E\left[u\right] \leq \gamma\left(x\right)$ we conclude that the second addend in (22) is negative if F is logconcave and $f'F/f^2$ is non-decreasing. For example, if F is a power distribution (with uniform as a special case), L/F is convex (being linear), F is logconcave, and $f'F/f^2$ is constant, so that γ is concave. Lemma B.1 extends this logic:

[65] These results follow from Prékopa's (1973) Theorem, which guarantees that logconcavity is preserved by integration; see, for example, An (1998) and Bagnoli and Bergstrom (2005).

Lemma B.1 *Distribution S crosses only once and from above distribution N in the interior of the support if:*
(i) the distribution F is logconcave,
(ii) the mean-advantage-over-inferiors L/F is convex, and
(iii) the ratio of the distribution to the density F/f is logconcave.

Proof of Lemma B.1. Note the following facts:

(a) By the law of iterated expectations S and N have the same expectation so that they must cross at least once in the interior of the common support, i.e., γ has at least one interior fixed point.

(b) Assumption (i), logconcavity of F, implies logconcavity of L by Prékopa's (1973) Theorem. Then, by the first result shown in the proof of Proposition B.1, at the last interior fixed point γ must cross the 45-degree line from below.

To claim that there is a single interior crossing between S and N it is sufficient to show that the domain of γ can be partitioned into two connected regions: a lower region $\underline{\mathcal{S}}$ in which γ can cross the 45-degree line only from below and a higher region $\overline{\mathcal{S}}$ in which γ can cross the 45-degree line only from above. At a fixed point x^*, γ crosses the 45-degree line from below if and only if

$$\left.\frac{d\gamma(x)}{dx}\right|_{x=x^*} = \underbrace{\frac{U'(x)}{2}}_{\leq 1 \text{ if } F \text{ is logconcave}} \cdot \frac{\frac{f(\mathcal{U}(x^*)-E[u])}{F(\mathcal{U}(x^*)-E[u])}}{\frac{f(\gamma(x^*))}{F(\gamma(x^*))}} > 1,$$

If the inequality is reversed, at x^*, γ crosses the 45-degree line from above.

To prove that $\underline{\mathcal{S}}$ and $\overline{\mathcal{S}}$ are connected and that $\overline{\mathcal{S}} \leq \underline{\mathcal{S}}$ in the set order (i.e., all elements of $\overline{\mathcal{S}}$ are lower than all elements of $\underline{\mathcal{S}}$) we show that under (i), (ii), (iii) there exists $\bar{y} \in [\mathcal{U}(\underline{u}+E[u]), \overline{u}]$ such that if y is fixed point of γ

$$\begin{cases} \frac{\frac{f(\mathcal{U}(x^*)-E[u])}{F(\mathcal{U}(x^*)-E[u])}}{\frac{f(x^*)}{F(x^*)}} > \frac{2}{\mathcal{U}'(x^*)} & \text{if } y \in [\mathcal{U}(\underline{u}+E[u]), \bar{y}] \\ \frac{\frac{f(\mathcal{U}(x^*)-E[u])}{F(\mathcal{U}(x^*)-E[u])}}{\frac{f(x^*)}{F(x^*)}} < \frac{2}{\mathcal{U}'(x^*)} & \text{if } y \in [\bar{y}, \overline{u}]. \end{cases} \tag{24}$$

Note that, from (6) the right-hand side $\frac{2}{\mathcal{U}'(x^*)}$ is increasing if (ii) $\frac{L}{F}$ convex. So it is sufficient to show that the left-hand side is decreasing. Indeed the ratio of reverse hazard rates in (24) is decreasing in x^* if

$$\frac{d}{dx}\frac{\frac{f(\mathcal{U}(x)-E[u])}{F(\mathcal{U}(x)-E[u])}}{\frac{f(x)}{F(x)}} = \frac{\frac{f'(\mathcal{U}(x)-E[u])F(\mathcal{U}(x)-E[u])-f(\mathcal{U}(x)-E[u])^2}{F(\mathcal{U}(x)-E[u])^2}\mathcal{U}'(x)\frac{f(x)}{F(x)} - \frac{f(\mathcal{U}(x)-E[u])}{F(\mathcal{U}(x)-E[u])}\frac{f'(x)F(x)-f(x)^2}{F(x)^2}}{\left(\frac{f(x)}{F(x)}\right)^2} < 0.$$

Rewriting this using logconcavity of F, (i), which implies $f'F - f^2 < 0$, we have

$$\frac{\frac{F(\mathcal{U}(x)-E[u])}{f(\mathcal{U}(x)-E[u])}\frac{f(\mathcal{U}(x)-E[u])^2-f'(\mathcal{U}(x)-E[u])F(\mathcal{U}(x)-E[u])}{F(\mathcal{U}(x)-E[u])^2}}{\frac{F(x)}{f(x)}\frac{f(x)^2-f'(x)F(x)}{F(x)^2}} = \frac{\frac{f(\mathcal{U}(x)-E[u])}{F(\mathcal{U}(x)-E[u])}-\frac{f'(\mathcal{U}(x)-E[u])}{f(\mathcal{U}(x)-E[u])}}{\frac{f(x)}{F(x)}-\frac{f'(x)}{f(x)}} > \frac{1}{\mathcal{U}'(x)}. \tag{25}$$

Recalling that $x \geq \mathcal{U}(x) - E[u]$ for all x, by logconcavity of $\frac{F}{f}$, (iii), we conclude that the left-hand side of (25) is larger than 1. Then (25) holds because $\frac{1}{\mathcal{U}'(x)} \leq 1$ for all x.

Because $\underline{\mathcal{S}}$ and $\overline{\mathcal{S}}$ are connected and $\overline{\mathcal{S}} \leq \underline{\mathcal{S}}$ under the set order, claim (b) implies that there cannot be a fixed point in the interior of $\overline{\mathcal{S}}$. To see this, suppose, by way of contradiction, that there is such a fixed point in the interior of $\overline{\mathcal{S}}$, call it $\tilde{x}$; then, given that at $\tilde{x}$, γ crosses the 45-degree line from above and that there is no further crossing from above at any $x > \tilde{x}$, all points $x > \tilde{x}$ in the domain would satisfy $\gamma(x) < x$ violating claim (b), thus reaching a contradiction.

Given that by claim (a) there must be an interior crossing and that crossings must alternate in sign, the set $\underline{\mathcal{S}}$ contains a unique interior crossing, while $\overline{\mathcal{S}}$ can contain at most one crossing, located at the upper bound of the support. If $\overline{u}$ is bounded there is exactly one fixed point in set $\overline{\mathcal{S}}$ at the upper bound, where $\frac{d\gamma(x)}{dx} < 1$. If, instead, $\overline{u}$ is unbounded there is no crossing in $\overline{\mathcal{S}}$. **Q.E.D.**

Condition (i) implies a global version of Proposition B.1's condition (A); it is relatively weak given that logconcavity is preserved under integration. Condition (ii) is a mild regularity condition, which is equivalent to concavity of the expected sum $\mathcal{U}(u_d)$ as a function of the selectively disclosed value; see (6). Condition (iii) means that F is logconcave relative to f (see Whitt, 1985); logconcavity of F/f is automatically satisfied if f is logconvex and F is logconcave. Under these sufficient conditions, the domain of the function γ can be partitioned into two connected regions, a lower region in which $\gamma(x)$ can cross the 45-degree line x only from below and a higher region in which $\gamma(x)$ can cross x only from above. This property, weaker than concavity of γ, implies clockwise rotation because by definition $\gamma(x) > x \Leftrightarrow S(\mathcal{U}(x)) < N(\mathcal{U}(x))$. Below we report a number of examples of distributions satisfying Lemma B.1's three sufficient conditions: power distributions (including uniform), gamma(α, β) as well as Weibull with shape $\alpha \leq 1$ (including exponential), extreme value Gumbel, and Pareto with $\alpha > 1$ (for which the expectation exists).

Next, we show that logconcavity of L—which is implied by logconcavity of F by Prékopa's (1973) Theorem—is necessary for S to be a clockwise rotation of N.

Lemma B.2 *Distribution S is a clockwise rotation of distribution N only if the left-hand integral L is logconcave at the upper bound $\overline{u}$.*

Proof of Lemma B.2. We show by contradiction that if left-hand integral L is logconvex at the upper bound, S cannot be a clockwise rotation of N. Note that, if S is a clockwise rotation of N, then for all U above the rotation point $S(U) \leq N(U)$. However, as shown

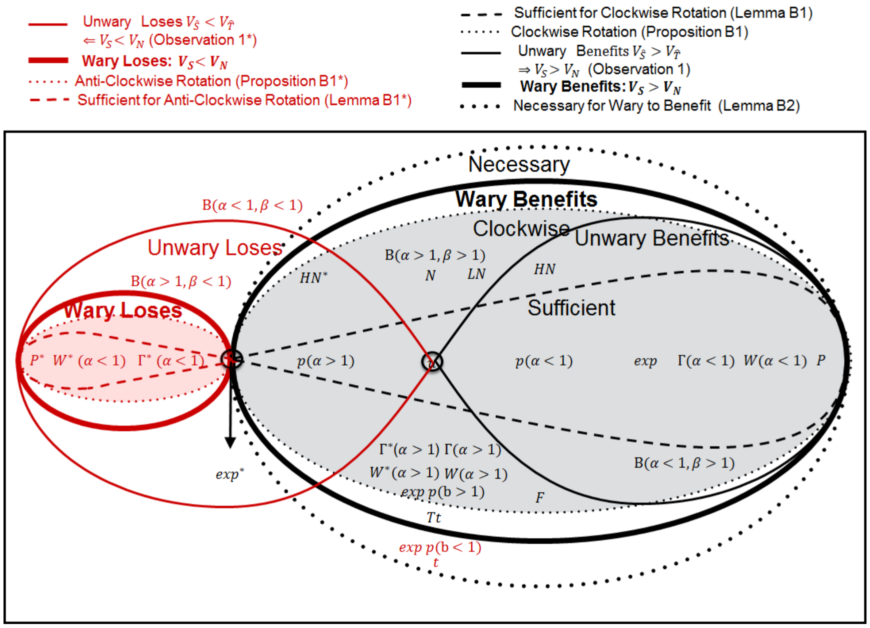

Figure 5: Summary of main results on impact of selective disclosure.

in the proof of Proposition B.1, if L is logconvex then $x > \gamma(x)$, or equivalently $S(U) > N(U)$, in a left neighborhood of the upper bound, reaching a contradiction. **Q.E.D**.

Reversing Proposition B.1's condition (A) we have:

Proposition B.1* *Distribution S crosses only once and from below distribution N in the interior of the support if and only if:*
(A) the left-hand integral $L(u) = \int_{\underline{u}}^{u} F(y)\,dy$ is logconvex at the upper bound and*
(B) the function γ has a unique fixed point in the interior of the domain $(\mathcal{U}^{-1}(\underline{u} + E[u]), \overline{u})$.

Similarly, turning Lemma B.1 on its head, the following result provides sufficient conditions for *anti-clockwise* rotation.

Lemma B.1* *Distribution S crosses only once and from below distribution N in the interior of the support if:*
(i) the distribution F is logconvex,*
(ii) the mean-advantage-over-inferiors L/F is convex, and
(iii) the ratio of the distribution to the density F/f is logconcave.

Proof of Lemma B.1*. The proof follows the same steps as for Lemma B.1 and reversing the sign of (25). **Q.E.D.**

Figure 5 summarizes the implications of selective disclosure for receiver welfare for a large set of parametric distributions (see below for verification):

N: normal (clockwise rotation)
LN: log-normal (clockwise rotation)
HN: half-normal ($V_{\hat{S}} > V_N$ and clockwise rotation)
$B(\alpha < 1, \beta > 1)$: beta ($V_{\hat{S}} > V_N$ and clockwise rotation)
$B(\alpha > 1, \beta > 1)$: beta (clockwise rotation)
$B(\alpha > 1, \beta < 1)$: beta ($V_{\hat{S}} < V_N$)
$B(\alpha < 1, \beta < 1)$: beta (first interior crossing from below, second from above)
$p\,(\alpha < 1)$: power ($V_{\hat{S}} > V_N$ and sufficient for clockwise rotation)
u: uniform ($V_{\hat{S}} \equiv V_N$ and sufficient for clockwise rotation)
G: extreme value Gumbel (sufficient for clockwise rotation)
$p\,(\alpha > 1)$: power ($V_{\hat{S}} < V_{\hat{T}}$ and sufficient for clockwise rotation)
$\Gamma\,(\alpha < 1)$: gamma ($V_{\hat{S}} > V_{\hat{T}}$ and sufficient for clockwise rotation)
$\Gamma\,(\alpha > 1)$: gamma ($V_{\hat{S}} > V_{\hat{T}}$ and clockwise rotation)
exp: negative exponential ($V_{\hat{S}} > V_{\hat{T}}$ and sufficient for clockwise rotation)
$W(a < 1)$: Weibull ($V_{\hat{S}} > V_{\hat{T}}$ and sufficient for clockwise rotation)
$W(a > 1)$: Weibull ($V_{\hat{S}} > V_{\hat{T}}$ and clockwise rotation)
P: Pareto ($V_{\hat{S}} > V_{\hat{T}}$ and sufficient for clockwise rotation)
t: Student's t (necessary for clockwise rotation)
F: Fisher-Snedecor F (clockwise rotation)
Tt: truncated Student's t ($V_S > V_T$)
HN^*: mirror half-normal ($V_{\hat{S}} < V_{\hat{T}}$ and clockwise rotation)
$\Gamma^*\,(\alpha > 1)$: mirror gamma ($V_{\hat{S}} < V_{\hat{T}}$ and clockwise rotation)
$\Gamma^*\,(\alpha < 1)$: mirror gamma ($V_{\hat{S}} < V_{\hat{T}}$ and anticlockwise rotation)
exp*: positive exponential ($V_S \equiv V_N$ and $S \equiv N$)
$W^*\,(\alpha < 1)$: mirror Weibull ($V_{\hat{S}} < V_{\hat{T}}$ and clockwise rotation)
P^*: mirror Pareto (sufficient for anti-clockwise rotation)

Sufficient Conditions for Rotation: Verification in Examples. The following inheritance property provides a set of sufficient conditions that will prove convenient to verify condition (ii) of Proposition B.1 in examples:

Lemma B.3 *(a) The ratio L/F is convex if (i) the ratio F/f is convex and (ii) $F\left(\underline{u}\right)^2/f\left(\underline{u}\right) = 0$ at the lower bound $\underline{u}$ of the support. (b) The ratio F/f is convex if (iii) the ratio f/f' is convex and (iv) $f\left(\underline{u}\right)^2/f'\left(\underline{u}\right) = 0$.*

Proof of Lemma B.3. The argument is similar to Baricz (2010), Theorem 2(c), and is based on the monotone form of de l'Hôpital rule; see Wu and Debanth (2009). Clearly, $L(u)/F(u)$ is convex whenever $L(u)\,f(u)/F(u)^2$ decreases in u. At the lower bound $\underline{u}$ of the support we have that $F\left(\underline{u}\right)^2/f\left(\underline{u}\right) = 0$ (by assumption) and $L\left(\underline{u}\right) = 0$ (by definition), so that

$$\frac{L(u)}{\frac{F(u)^2}{f(u)}} = \frac{L(u) + L\left(\underline{u}\right)}{\frac{F(u)^2}{f(u)} + \frac{F\left(\underline{u}\right)^2}{f\left(\underline{u}\right)}}$$

which, in view of the monotone form of the l'Hôpital rule, is decreasing if the function

$$\frac{\frac{dL(u)}{du}}{\frac{d\left(\frac{F(u)^2}{f(u)}\right)}{du}} = \frac{F(u)}{\frac{2f(u)^2F(u)-F(u)^2f'(u)}{f(u)^2}} = \frac{1}{2-\frac{F(u)f'(u)}{f(u)^2}}$$

is decreasing in u, proving part (i). The proof of part (ii) follows the same steps. **Q.E.D.**

We now derive a further sufficient condition for the mean-advantage over inferiors to be convex:

Lemma B.4 *L/F is convex if f is logconcave and decreasing.*

Proof of Lemma B.4. The assumptions imply that F/f is convex,

$$\frac{d^2}{du^2}\frac{F(u)}{f(u)} = \frac{-\left(f(u)f''(u)-f'(u)^2\right)F(u)+\left(F(u)f'(u)-f(u)^2\right)f'(u)}{f(u)^3} \geq 0$$

given that logconcavity of f (i.e., $ff''-f'^2 \leq 0$) guarantees that F is logconcave ($Ff'-f^2 \leq 0$) by Prepoka's Theorem. The result follows by Lemma B.3 given that we then have that $f(\underline{u}) > 0$ and thus $F(\underline{u})^2/f(\underline{u}) = 0$. **Q.E.D.**

Example B.2 (Power) *The power distribution with $f(u) = \alpha u^{\alpha-1}$ and support $(0,1)$ satisfies Lemma B.1. The uniform distribution corresponds to $\alpha = 1$.*

Verification of Example B.2. (i) F is logconcave, see Bagnoli and Bergstrom (2005).
(ii) $d(L(u)/F(u))/du = 1/(\alpha+1)$ so that L/F is linear.
(iii) $d^2[\ln(F(u)/f(u))]/du^2 = -\alpha/u^2 + (\alpha-1)/u^2 = -1/u^2 < 0$.

Example B.3 (Gamma) *The gamma(α,β) distribution with $\alpha \leq 1$ satisfies Lemma B.1.*

Verification of Example B.3. (i) F is logconcave, see Bagnoli and Bergstrom (2005).
(ii) L/F is convex by Lemma B.3.a given that

$$\lim_{x\to 0^+}\frac{F(x)^2}{f(x)} = 0,$$

as $\lim_{x\to 0^+} F(x) = 0$ and $\lim_{x\to 0^+} f(x) = +\infty$ for $\alpha < 1$, and that F/f is convex by Lemma B.3.b because

$$\frac{\partial^2}{\partial x^2}\frac{f(x)}{f'(x)} = \frac{\partial^2}{\partial x^2}\left(\frac{x\beta}{\beta(\alpha-1)-x}\right) = \frac{2\beta^2(\alpha-1)}{[\beta(\alpha-1)-x]^3} > 0$$

and

$$\lim_{x\to 0^+}\frac{f(x)^2}{f'(x)} = \lim_{x\to 0^+}\frac{x^\alpha e^{-\frac{x}{\beta}}}{\Gamma(\alpha)\beta^{\alpha-1}[\beta(\alpha-1)-x]} = 0.$$

(iii) F/f is logconcave because F is logconcave and f is logconvex for $\alpha \leq 1$.

Example B.4 (Gumbel) *The extreme value Gumbel distribution satisfies Lemma B.1.*

Verification of Example B.4. (i) F is logconcave.
(ii) L/F is convex by verifying the conditions of Lemma B.3.a.
(iii) F/f is loglinear.

Example B.5 (Pareto) *The Pareto distribution* $F(u) = 1 - \underline{u}^{\alpha} u^{-\alpha}$ *with* $\underline{u} > 0$ *and* $\alpha > 1$ *(so that the expectation exists) satisfies Lemma B.1.*

Verification of Example B.5. (i) F is logconcave, see Bagnoli and Bergstrom (2005).
(ii) By Lemma B.3.a, L/F is convex because (i) F/f is convex, given that $d(F/f)/du = [(1+a)u^{a} - \underline{u}^{\alpha}]/(\alpha \underline{u}^{\alpha})$ is clearly increasing in u, and (ii) $F(\underline{u})^2/f(\underline{u}) = (1 - \underline{u}^{\alpha}\underline{u}^{-\alpha})^2/(\alpha \underline{u}^{\alpha} \underline{u}^{-1-\alpha}) = 0$.
(iii) F/f is logconcave for $\alpha > 1$ because

$$\frac{d^2}{du^2} \ln \frac{F}{f} = -\frac{\overbrace{1 + \alpha - 2\left(\underline{u}/u\right)^{\alpha}}^{>0} + \overbrace{\alpha^2 \left(\underline{u}/u\right)^{\alpha} - \alpha \left(\underline{u}/u\right)^{\alpha}}^{>0} + \left(\underline{u}/u\right)^{2\alpha}}{u^2 \left[\left(\underline{u}/u\right)^{\alpha} - 1\right]^2} < 0.$$

Failure of Clockwise Rotation. We now characterize the set of *non-logconcave* distributions for which the clockwise rotation property does not hold.

Example B.6 (Positive Exponential) *Selective and non-selective disclosure induce identical distributions* $S(U) \equiv N(U)$ *in the boundary case with loglinear left-hand integral* L*, which corresponds to the positive exponential distribution with* $f(u) = \lambda e^{\lambda(u - \overline{u})}$ *with* $\lambda > 0$ *and support* $(-\infty, \overline{u}]$*.*

Verification of Example B.6. From $F(u) = e^{\lambda(u-\overline{u})}$ and $E[u] = \overline{u} - 1/\lambda$ we have $N(U) = F(U - E[u]) = e^{\lambda\left(U - \left(\overline{u} - \frac{1}{\lambda}\right) - \overline{u}\right)} = e^{U\lambda - 2\overline{u}\lambda + 1}$. From $U = E[u^1 + u^2 | u_d = \max\langle u^1, u^2 \rangle] = 2u_d - 1/\lambda$ we have $u_d(U) = (U + 1/\lambda)/2$, so that $S(U) = (F(u_d(U)))^2 = e^{2\lambda\left(\frac{U + \frac{1}{\lambda}}{2} - \overline{u}\right)} = e^{U\lambda - 2\overline{u}\lambda + 1}$.

The rotation property is reversed for the mirror image of the Pareto distribution, a logconvex distribution that has thicker tail than the positive exponential distribution on both sides:

Example B.7 (Mirror Pareto) *The mirror image of the Pareto distribution with* $f(u) = -(\beta/\overline{u})(u/\overline{u})^{-\beta}$ *and support* $(-\infty, \overline{u}]$ *with* $\overline{u} < 0$ *and* $\beta > 1$ *(so that the expectation exists) satisfies all three conditions of Lemma B.1*.*

Verification of Example B.7. We have $E[u] = -\beta/(\beta-1)$, $F(u) = (u/\overline{u})^{-\beta}$, $L(u) = -u/(1-\beta)(u/\overline{u})^{-\beta}$ and $E[u|u < u_d] = \beta u_d/(\beta-1)$, so that:
(i*) F is logconvex: $d^2 \ln F(u)/du^2 = \beta/u^2 > 0$.
(ii) L/F is concave (being linear): $d^2(L(u)/F(u))/du^2 = d(-u/(1-\beta))/du^2 = 0$.
(iii) F/f is logconcave: $d^2 \ln(F(u)/f(u))/du^2 = \beta/u^2 - (1+\beta)/u^2 = -1/u^2 < 0$.

We conclude by illustrating two classes of examples for which violation of logconcavity of L result in S and N that cross more than once in the interior. First, we give an example of a distribution with U-shaped density that is steeper than the positive exponential at the top of the support:

Example B.8 (Beta) *If F is beta(α, β) with parameter $\beta < 1$, distributions S and N cross an even number of times, so that $V_S(R) < V_N(R)$ for sufficiently high R and $V_S(R) > V_N(R)$ for sufficiently low R.*

Verification of Example B.8. Logconcavity of the left-hand integral L is violated for u sufficiently close to $\overline{u} = 1$ because then $\lim_{u\to 1} L(u) f(u) - F(u)^2 = \infty$ given that $\lim_{u\to 1} L(u) = 1 - \mu$ and $\lim_{u\to 1} f(u) = \lim_{u\to 1} u^{\alpha-1}(1-u)^{\beta-1} = \infty$ for $\beta < 1$. Thus, by Lemma B.2 at the last crossing S crosses N from below. Given that $\underline{u} > -\infty$, at the first crossing S crosses N from above. We conclude that S crosses N an even number of times.

Second, distributions F with a bottom tail that is thicker than the negative exponential (thus violating of logconcavity of L) result in S first crossing N from below:

Example B.9 (Exponential Power) *If F is exponential power with shape parameter $b \geq 1$ (including as special case Laplace for $b = 1$, Normal for $b = 2$, and uniform for $b \to \infty$), S crosses N once and from above in the interior. If $b < 1$ (so that the bottom tail is are thicker tail than negative exponential), distribution S first crosses N from below and then a second time from above.*

Example B.10 (Student's t) *If F is Student's t with at least 2 degrees of freedom (so that the expectation exists), distribution S first crosses N from below and then a second time from above.*

Example B.11 (Truncated Student's t) *If F is a left truncation of Student's t with sufficiently large variance, S crosses N three times (from above, below, and above); if an addition the variance is not too large, so that the first and second crossing points are sufficiently close, the welfare property $V_S > V_N$ is preserved in spite of violation of the clockwise rotation Property 2.*

B.3 Different Weights for Different Attributes

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

$$u = \alpha^1 u^1 + \alpha^2 u^2,$$

where the weights satisfy without loss of generality $\alpha^1 > \alpha^2 > 0$. We focus on the tractable case where u^1 and u^2 are independent and uniformly distributed on $[\underline{u}, \overline{u}]$ and 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, 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

$$u^1 \geq \frac{(\alpha^1 - \alpha^2)}{\alpha^1}\underline{u} + \frac{\alpha^2}{\alpha^1}u^2. \tag{26}$$

If (26) does not hold, $d = 2$ is disclosed. With this rule at hand, after disclosing $d = 1$ the expected utility equals

$$U = \frac{3}{2}\alpha^1 u^1 - \frac{1}{2}\left(\alpha^1 - 2\alpha^2\right)\underline{u},$$

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

$$U = \frac{3}{2}\alpha^2 u^2 + \frac{1}{2}\left(2\alpha^1 - \alpha^2\right)\underline{u},$$

so that now $U \in \left[(\alpha^1 + \alpha^2)\underline{u}, \frac{3}{2}\alpha^2\overline{u} + \frac{1}{2}\left(2\alpha^1 - \alpha^2\right)\underline{u}\right]$. Note that from $\alpha^1 \geq \alpha^2$, which we stipulated without loss of generality, the highest value of U is attained when disclosing $u^1 = \overline{u}$, $\overline{U} = \frac{3}{2}\alpha^1\overline{u} - \frac{1}{2}\left(\alpha^1 - 2\alpha^2\right)\underline{u}$, while $\underline{U} = (\alpha^1 + \alpha^2)\underline{u}$. After some calculations we obtain the following characterization

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

where

$$U' = \frac{3}{2}\alpha^2\overline{u} + \frac{1}{2}\left(2\alpha^1 - \alpha^2\right)\underline{u},$$

with $U' \in \left(\underline{U}, \overline{U}\right)$ for $\alpha^1 > \alpha^2$.

We next derive $\hat{S}(U)$ (where the receiver is unwary 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

$$u^1 \geq \frac{(\alpha^1 - \alpha^2)}{\alpha^1} \frac{(\overline{u} + \underline{u})}{2} + \frac{\alpha^2}{\alpha^1} u^2,$$

and otherwise to disclose $d = 2$. Again after some calculations we obtain

$$\hat{S}(U) = \begin{cases} \frac{\left[2U - (\alpha^1+\alpha^2)\left(\frac{(\overline{u}+\underline{u})}{2}+\underline{u}\right)\right]^2 - \left[(\alpha^1-\alpha^2)\left(\frac{(\overline{u}+\underline{u})}{2}-\underline{u}\right)\right]^2}{4\alpha^1\alpha^2(\overline{u}-\underline{u})^2} & \text{for } \underline{\hat{U}} \leq U \leq \hat{U}' \\ \frac{2U - \alpha^2(\overline{u}+\underline{u}) - 2\alpha^1\underline{u}}{2\alpha^1(\overline{u}-\underline{u})} & \text{for } \hat{U}' < U \leq \overline{\hat{U}} \end{cases}$$

where

$$\begin{aligned} \underline{\hat{U}} &= \alpha^2\underline{u} + \alpha^1 \frac{(\overline{u}+\underline{u})}{2}, \\ \overline{\hat{U}} &= \alpha^1\overline{u} + \alpha^2 \frac{(\overline{u}+\underline{u})}{2}, \\ \hat{U}' &= \alpha^2\overline{u} + \alpha^1 \frac{(\overline{u}+\underline{u})}{2}, \end{aligned}$$

with $\hat{U}' \in \left(\underline{\hat{U}}, \overline{\hat{U}}\right)$ for $\alpha^1 > \alpha^2$.

What now complicates the analysis is that with unequal weights $\alpha^1 \neq \alpha^2$ the sender strategy is no longer immediate. Without loss of generality, we can limit the sender 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

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

for $U \in \left[\alpha^1\underline{u} + \alpha^2\frac{(\overline{u}+\underline{u})}{2}, \alpha^1\overline{u} + \alpha^2\frac{(\overline{u}+\underline{u})}{2}\right]$. When the sender discloses $d = 2$, then

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

for $U \in \left[\alpha^2\underline{u} + \alpha^1\frac{(\overline{u}+\underline{u})}{2}, \alpha^2\overline{u} + \alpha^1\frac{(\overline{u}+\underline{u})}{2}\right]$. Overall, we have:

Proposition B.2 *When u^i is uniformly distributed but the receiver applies different weights, $\alpha^1 \neq \alpha^2$, the distributions for the receiver's perceived utility are related by FOSD (unwary) and Clockwise Rotation (wary).*

B.4 Selective Non-Disclosure

In this Appendix we formalize the selective non-disclosure setting sketched in the main text, thereby establishing the robustness of the impact of selective disclosure on both the unwary and the wary receiver (as captured in Properties 1 and 2 in the main text) for an alternative underlying disclosure model.

Recall that the selective non-disclosure model seeks to capture situations in which the sender decides whether or not to disclose information at all, depending on what the sender has learned about the offering's value to the receiver. Denote the receiver's *true* utility from the sender's offering by u which is distributed according to $F(u)$ admitting an atomless density $f(u)$, with support $[\underline{u}, \overline{u}]$. Denote the expectation of u by $E[u]$. The sender learns about u with probability θ and then decides whether or not to disclose u to the receiver.[66] The concrete value of $\theta \in \{\theta^L, \theta^H\}$, $0 < \theta^L < \theta^H < 1$, i.e., the likelihood with which the sender is informed, will depend on the sender's information acquisition strategy. As long as $\theta < 1$, the fact that the sender remains ignorant with strictly positive probability prevents full unravelling of the information about receiver preferences that the sender has actually learnt. As is well known from, e.g., Jung and Kwon (1988), the sender who wants to maximize the receiver's perceived valuation will apply a threshold rule: u is disclosed only when it is above a threshold $u_d(\theta)$, at which the receiver's perceived valuation without disclosure is just equal to the disclosed true utility. That is, for given θ, the sender only discloses when $u \geq u_d(\theta)$, which uniquely solves

$$u_d(\theta) = \frac{(1-\theta)E[u] + \theta F(u_d(\theta))E[u \mid u \leq u_d(\theta)]}{(1-\theta) + \theta F(u_d(\theta))}. \tag{27}$$

Existence of a unique solution $u_d(\theta)$ follows immediately, as, using integration by parts, (27) can be rewritten to obtain

$$(1-\theta)\left[E[u] - u_d(\theta)\right] - \theta \int_{\underline{u}}^{u_d(\theta)} F(u)du = 0, \tag{28}$$

where the left-hand side is strictly positive for $u_d = \underline{u}$, strictly negative for $u_d > E[u]$ and strictly decreasing in u_d. At $u_d(\theta)$ 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 (27).[67]

The main object of our analysis in the main text is the *ex-ante* distribution of the receiver's perceived utility from the sender's offering U which we denote by $G(U)$. For a given and known value of θ, this distribution is $G(U) = (1-\theta) + \theta F(U)$ for $u_d(\theta) \leq U \leq \overline{u}$, with a mass point at the lower bound.[68] We now evaluate how this distribution of the receiver's perceived utility changes when the sender is more likely to be informed as he has acquired information about receiver preferences. I.e., we compare $G(U)$ for two different values $0 < \theta^L < \theta^H < 1$, where we denote $G(U) = N(U)$ for $\theta = \theta^L$ and $G(U) = S(U)$ for $\theta = \theta^H$. The left-side panel of Figure 6 depicts this comparison for the case of a standard uniform distribution of u. The key observation is that when the sender is more likely to be informed about the receiver's preferences, the respective distribution (that is, for θ^H) is obtained through a clockwise rotation from the distribution for the lower value θ^L such

[66] When the sender remains uninformed, we may stipulate that he either cannot disclose the respective information or it is not optimal to do so. Else, there is simply no scope for *selective* non-disclosure.

[67] E.g., for the standard uniform distribution we get $u_d(\theta) = \sqrt{1-\theta}\left(1 - \sqrt{1-\theta}\right)/\theta$.

[68] E.g., for the standard uniform distribution we have $G(U) = (1-\theta) + \theta U$ for $u_d(\theta) \leq U \leq 1$ (with an atom at $U = u_d(\theta)$).

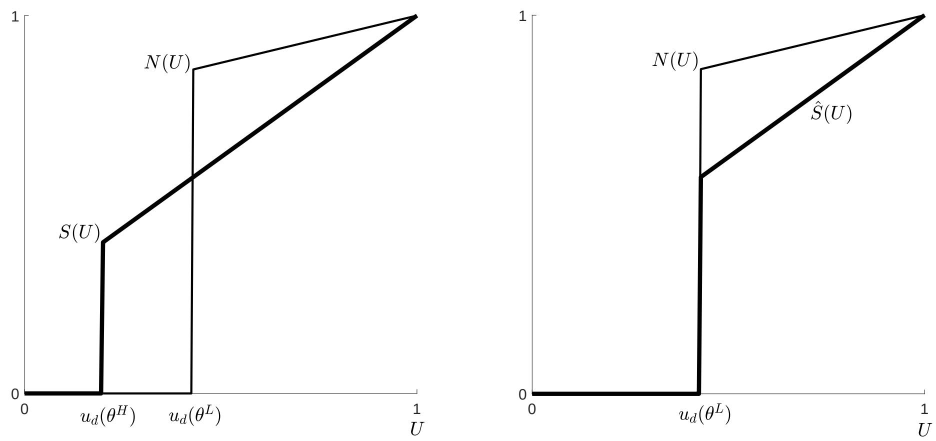

Figure 6: **Comparison of ex-ante distributions of the receiver's perceived utility for the selective non-disclosure model.** The left panel shows the transition from less selective ($\theta^L = \frac{1}{4}$, thin line) to more selective disclosure ($\theta^H = \frac{3}{4}$, bold line) with a wary receiver and the right panels the respective transition with an unwary receiver for the case of a standard uniform distribution of u.

that Property 2 in the main text holds also for this disclosure model. That is, increased selectivity of disclosure (by a better informed sender) moves probability mass away from the center and into the tails of the distribution. This is intuitive as, first, facing a better informed sender, a receiver adjusts downwards his perceived valuation without disclosure, leading to a lower threshold $u_d(\theta)$ and, thus, a larger set of disclosed values and, second, as disclosure requires the sender to be informed, also the probability of disclosure increases for any $u \geq u_d(\theta)$. The following Proposition shows the robustness of this intuition and establishes the general validity of Property 2 in the selective non-disclosure model:

Proposition B.3 *Consider the setting of selective non-disclosure. If it is commonly known that the probability the sender is informed about the preferences of the receiver increases from θ^L to $\theta^H > \theta^L$, then the resulting shift in the ex-ante distribution of the receiver's perceived utility U from $G(U) = N(U)$ to $G(U) = S(U)$ represents a clockwise rotation as defined in Property 2.*

Proof of Proposition B.3. Note, first, that we have from (28) that

$$\frac{du_d(\theta)}{d\theta} = -\frac{[E[u] - u_d(\theta)] + \int_{\underline{u}}^{u_d(\theta)} F(u)du}{(1-\theta) + \theta F(u_d(\theta))} < 0. \tag{29}$$

To characterize $N(U)$ and $S(U)$, where θ 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

$$\frac{dj(\theta)}{d\theta} = \theta \frac{du_d(\theta)}{d\theta} f(u_d(\theta)) - (1 - F(u_d(\theta)) < 0, \tag{30}$$

where we made use of (29). As we move from $\theta = \theta^L$ to $\theta = \theta^H$, Property 2, thus, follows directly from (29), (30), and the fact that $G(U)$ is strictly decreasing in θ (with slope $F(U) - 1$) for all $U > u_d(\theta)$. Note also that the single crossing point is at $u_d(\theta^L)$, the lower support of $N(U)$. **Q.E.D.**

So far we presumed that the receiver is aware of the fact that the sender is more likely to be informed about her preferences (with $\theta = \theta^H$). Suppose now instead that the receiver wrongly believes that the sender only observes her preferences with the strictly lower probability θ^L and denote the implied distribution of the receiver's perceived utility by $G(U) = \hat{S}(U)$. Note that then the sender optimally still applies the cutoff $u_d(\theta^L)$, but now, when there is no disclosure, the receiver's *perceived* value U differs from the *true* conditional expected value.[69] The right-side panel of Figure 6 shows that, compared to the case where indeed $\theta = \theta^L$, the sender's non-anticipated higher likelihood of observing the receiver's preferences results in a FOSD shift of the ex-ante distribution of U. Put differently, 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. The following Proposition formalizes these results:

Proposition B.4 *Consider the setting of selective non-disclosure. If the receiver is unaware that the sender is more likely to be informed, then Property 1 holds, i.e., the distribution of the receiver's perceived utility $G(U) = \hat{S}(U)$ dominates $N(U)$ in the sense of first-order stochastic dominance.*

Proof of Proposition B.4. When the receiver is unwary of the switch to $\theta = \theta^H$, we have that the distribution $\hat{S}(U)$ is zero up to $u_d(\theta^L)$, then jumps to $1 - \theta^H \left(1 - F(u_d(\theta^L))\right)$, and is equal to $\left(1 - \theta^H\right) + \theta^H F(U)$ for $[u_d(\theta^L), \overline{u}]$. Thus, there is a FOSD shift, given that $\hat{S}(U) < N(U)$ holds strictly for all $u \in [u_d(\theta^L), \overline{u})$, i.e., at all points of the joint support of S and N apart from the upper bound.[70] **Q.E.D.**

[69] The concrete difference is $\left(\theta^H - \theta^L\right) F(u_d(\theta^L)) \left\{E[u] - E[u \mid u \leq u_d(\theta^L)]\right\} / \left[S(u_d(\theta^L))N(u_d(\theta^L))\right]$.

[70] All the results extend also to the case with an unbounded upper support $\overline{u} \to \infty$.