Markets for leaked information[*]

Steffen Huck   Georg Weizsäcker

October 27, 2015

Abstract

We study markets for sensitive personal information. An agent wants to communicate with another party but any revealed information can be intercepted and sold to a third party whose reaction harms the agent. The market for information induces an adverse sorting effect, allocating the information to those types of third parties who harm the agent most. In equilibrium, this limits information transmission by the agent, but never fully deters it. We also consider agents who naively provide information to the market. Their presence renders traded information more valuable and, thus, harms sophisticated agents by increasing the third party's demand for information. Half-baked regulatory interventions may hurt naive agents without helping sophisticated agents. Comparing monopoly and oligopoly markets, we find that oligopoly is often better for the agent: it requires a higher value of traded information and therefore has to grant the agent more privacy.

1 Introduction

Some information travels far and often in ways that its originators fail to foresee. From their perspective, the information "leaks"; alternatively, an outside observer might call the senders "naive" for providing the information in the first place without taking care to communicate safely. New developments in information technology breed such naivete. Few people keep pace with the state of the art of data collection, data storage and data processing. Even fewer can anticipate what a future agent controlling some future technology will do with information that is recorded today. Yet many households appear not to worry and reveal sensitive information, especially in online social media and online commerce, without taking precautionary measures.[1]

This paper investigates the market incentives arising from the possibility of leaked (or intercepted) personal information. We consider a player A who has some information to share but faces a trade-off. On the one hand, she wants to share her information with player B because of some joint activity that benefits from truthful revelation. On the other hand, A fears the information ending up in the wrong hands—a player C who may take actions against A's interest if he gets hold of her information. Such a trade-off in information revelation features in many economic applications. For example, a firm may want to share details of their technology with a supplier yet fears that competitors may get access to the information. Consumers may want to reveal their preferences at one shop yet prefer that the information is not shared with other shops to rule out price discrimination. Or, in a broader social realm, friends might want to exchange information about political, religious or sexual preferences yet may have reason to fear that this information is obtained by their employers, relatives or certain state agencies.

An important question is via what mechanism A's information can end up in C's hands. Our model studies the case where the information is *traded* in a market. There are two separate ways in which A's data may be available for trade that our model captures. First, player B might engage in double dealing using the information fruitfully for his mutual business with A, but also selling it off to C for additional private gain. Second, and analogously, A's information may be intercepted by another player (not B himself) who

---

[1]See the evidence reported in Acquisti and Grossklags, 2005, Norberg et al., 2007, Tsai et al., 2011, and Beresford et al., 2012, as well as the survey article by Acquisti et al., 2014.

sells it off to C.[2]

Depending on the anticipated market allocation that arises from B's and C's interaction, Player A must be careful when communicating with B. A key variable for A to anticipate is the cost that B incurs for intercepting and selling the information. He will offer A's data to C if its expected market revenue weakly exceeds the investment cost required for data processing and marketing, including storage, formatting, refinement, market transaction cost etc. In short, the cost reflects the technology that is available at the time of B's market entry decision. We argue that this technology is difficult for A to assess. The typical consumer and user of online media may vastly underestimate the ease with which her data are captured, screened and evaluated. Even institutional users like firms or other (non-IT) professionals may have little information about the exact routing of their emails and the security of their data.[3] A further reason for such biases is that the market entry decision may lie in the future, potentially years after the information revelation. Over time, advances in information technology drive down the cost of information processing and some of these advances are hidden from the affected users (and with few incentives for the informed parties to reveal them).[4]

We capture these misperceptions by assuming that there are two cognitive types of player A: sophisticated types who know the true cost and naive types of player A who *over-estimate* the cost of selling information. In all other respects, our model assumes agents to be fully rational. In order to allow for such partial naivete, we formulate a game with non-common priors where both types of A agents are aware that there is another type with different beliefs about the cost of information provision. In an extension, we also study a different type of naivete, where some agents do not fully understand the market allocation.

Our main analysis considers the case of a monopolistic market structure

---

[2]For simplicity, we model this situation also as if it was B himself who sold the information on the market. As there are no complementarities or other spillovers between the two actions—using the information for generating joint surplus with A and selling it off—B's utility will be separable in these two actions and can, thus, also capture the behavior of two different agents.

[3]Prominent recent cases of commercial data security breaches include Sony's hack in 2014 and Ashley Madison's hack in 2015.

[4]If the seller is different from the recipient of A's information, there are additional costs of obtaining the information, be it through interception or hacking.

(Section 3.2) where player A reveals her information only to one, possibly two-faced, player B.[5] Player B offers to C the sale of A's information at a monopolistic price. If C obtains the information, he reacts with an action that suits himself and harms A.C's reaction depends on the knowledge of his own type and of A's preference type. Different types of C are more or less aggressive, but all types of C need to obtain information about A in order to identify their optimal reaction. C's preference for obtaining the information are unknown to A and B, with a distribution that generates a standard market demand curve.

Our first result bears bad news for A: the monopolist sells the information to those types of C who harm A the most. This result follows from the assumption on opposite preference alignment of A and C. C's willingness to pay for the information increases monotonically with the harm that he inflicts on A. The result is, of course, a rather generic property of markets: they allocate goods to those buyers with the highest willingness to pay— here, the most aggressive C-types. But the harmful effect on A comes with a self-moderating equilibrium effect. Player A reacts to the danger of leakage by withholding information. The market traders B and C, in turn, are aware of A's reluctance and react to it by exercising caution themselves: market activity will never be so intense as to completely shut player A up.

The market's reaction to A's withholding of information also highlights how differently the market traders benefit from trading information that originates from naive versus from sophisticated As. Sophisticated A correctly anticipates the market activity and withholds information whenever the equilibrium predicts the possibility of information leakage. The market traders therefore cannot earn much profit from her. Naive A, in contrast, is an easy victim and generates larger profits. For many parameter constellations, she provides strictly more information than sophisticated A, and the parameter constellation where B's profit from trading information is maximal is one where naive A feels uninhibited in her information provision while sophisticated A withholds information maximally.

We also find an unambiguous effect of A's possible naivete on her own payoff. Unsurprisingly, naive A receives smaller payoffs than sophisticated A. But naive A also exerts a negative externality on sophisticated A. This

---

[5]In the alternative interpretation of the model, where B's information is intercepted, this corresponds to the case there is only one agent who can spy on A and sell the information.

externality arises because the prevalence of naivete increases the value of information, hence increases demand, and hence increases trading activity. Sophisticated A therefore needs to reduce information provision further, in response to naive A's presence. Ironically, a larger proportion of naive players may benefit a given naive player. As she realizes that players C can expect more information transmission with more players A who are like herself, she might become more cautious.

One possible consumer protection intervention is to educate consumers about the cost of obtaining and selling information. If such measures reduce the share of naive types, sophisticated As will benefit from them, while naive types who remain obstinate may be worse off. Other possible regulatory interventions affect B's cost of supplying information, e.g. through data security requirements or through bans of certain technologies. Our analysis shows how in equilibrium, such interventions may help A to provide information to B. We identify sufficient conditions under which all types of player A are better off if the cost of selling the information increases: if data protection is initially weak then a strong enough intervention helps player A for sure. However, we also show that weak or ill-designed interventions—even if they do increase B's costs—can cause harm without creating any benefit: they may encourage carelessness of naive types of A, without benefitting any of the sophisticated types.

Section 4 contains a number of extensions and robustness checks. Among them is the case of Bertrand oligopoly, where we allow for multiple entrants in the market. For player A, a competitive pressure in the market for information holds the promise that the information may not be traded at all because the equilibrium profits are small in expectation for any individual potential entrant. But a competitive market structure also implies the danger that more than one supplier of information operates in the market, which drives prices down and thus increases the frequency with which information is sold. We show that A's incentives to provide information are decisive for welfare: in equilibrium of the oligopoly market, the quality of the traded information must be higher than in the corresponding monopoly market (with identical entry cost per firm). Therefore, the oligopoly market structure must induce A to provide more precise information than monopoly, which, in turn, implies that sophisticated A prefers oligopoly over monopoly.[6] In other words, the competitive pressure forces the oligopolists to treat A well and grant her

---

[6]In one parameter region, it may be that naive A prefers monopoly.

more data security.

In Section 2 we discuss possible applications of our model in more detail and relate them to existing strands of the literature. Sections 3 and 4 contain the formal analysis and Section 5 concludes.

2 Applications and literature

Business-to-business interaction. Player A is a manufacturer who requires some specific inputs such as software or machinery for which she has to share detailed information about her technology with supplier B. This information is valuable to A's competitors who might use it for their strategic advantage or to copy A's technology. Similar problems arise with outsourcing such that, as Arrunada and Vazquez (2006) discuss in contexts of electronics and car manufacturing, A's contract manufacturer can become her competitor. Ho (2009) has a related theoretical discussion on outsourcing showing that the presence of multiple upstream trading partners can alleviate the contracting problems that arise from the threat of information leakage.

Social media and the work place. Player A is an employee who uses online social media (OSM) services with a friend or colleague, player B. In their conversations, A directly benefits from sharing details about her workplace situation with B. But her communication is recorded and may be offered on a market by a data aggregator who sieves A's OSM usage data. The potential buyer, player C, is player A's current or future employer who cares about A's focus on her job or her views about the firm's management. If A's communication reveals some deviations from the employer's expectations he can as the current employer redesign A's job, redeploy or sack her. Future potential employers may think carefully about what type of job to offer to her (if any). Related evidence is in Acquisti and Fong's (2014) large-scale field experiment showing some US employers indeed rely on OSM usage and that job market prospects are significantly affected through online revelation of personal traits.

Personalized advertising and pricing. Player A is a consumer whose activities are recorded through store cards, cookies, or other software that tracks her activities. Player B is a seller who caters to A's demands (generating surplus for both of them) but also collects the information about A's type which he can sell to other companies.

There are fast-growing theoretical and empirical literatures that investigate personalized advertisement and personalized pricing. For the theoretical literature see e.g. Taylor (2004), Armstrong and Zhou (2010), Jentzsch (2014), Bergemann and Bonatti (2015) and the survey on personalized pricing by the UK's Office of Fair Trading (2013). For empirical evidence on firm behavior see e.g. the field experiment by Vissers et al (2014). In these literatures, it is a widespread view that consumers may not be aware of the possibility of subsequent data use and the allocation of information happens via market mechanisms. Evidence on consumer behavior that is (differentially) consistent with several kinds of consumer naivete is discussed in the experiments and surveys in Acquisti and Grossklags (2004), Norberg et al, (2007), Tsai et al (2011), Beresford et al (2012) and Schudy and Utikal (2014). A recent example of an ill-understood market mechanism is consumer data provision by mobile phone applications to data aggregators. Very few telephone users ask themselves why free-of-charge software is provided in their Appstore. The answer they would get (if they cared to ask) is that these apps routinely transfer personalized data from the telephone to data aggregators, usually for advertisement purposes, and usually in return for money.

We point out to the reader that the cited theoretical papers do not describe special cases of our model. Our assumptions differ from those made in the theoretical literature on consumer privacy in at least two basic features. First, we focus on agents who only provide information and hence have no bargaining power and cannot achieve good deals in return for their information provision. Their only weapon is to withhold information—with the drawback of reducing the benefit from interacting with player B. Second, we focus on the case where the buyer of the information, player C, has incentives that are known to be oppositely aligned to those of player A. That is, his reaction can only harm A and the uncertainty about C pertains merely to the strength of his preference or ability to harm A.

Finally, the literature on contracting with differentially naive consumers (e.g. Gabaix and Laibson, 2006, Eliaz and Spiegler, 2006, Koszegi, 2014) has related assumptions on consumer naivete, and show a large number of other incentive effects that the presence of naive consumers entails and that we ignore in our analysis. For example, Heidhues and Koszegi (2014) point out the dynamic effects that arise if a seller can screen the naives from the sophisticates. Armstrong (2015) gives a comprehensive overview of search in markets with naive agents. Differently to large parts of this literature,

we find that sophisticated agents do not benefit from others being naive but instead they unambiguously suffer from having naive counterparts.

3 The model

3.1 Main elements

We start the exposition with a summary of the model's main architecture, leaving the details to the subsequent subsections.

Player A has private information $\theta_A \in \mathbb{R}$ that she wants to reveal to player B but not to player C. A sends a message m_A to B, which is constructed such that when B reads it, he either learns the true θ_A or receives an empty message. Player A's only choice variable in our model is the precision of her message to B: she picks the probability x_A with which B learns θ_A from her message. Increasing the probability x_A is costly and we assume that the cost $c_A(x_A) \in \mathbb{R}$ is increasing and convex.[7]

Player B makes two decisions. First, he picks an action $d_B \in \mathbb{R}$ for his joint project with player A. The action aims to match his beliefs about θ_A and the more precisely it is matched to the truth, the more surplus it generates. Second, B can offer to player C to sell him the message m_A. This offer causes cost c_B to player B.

If B decides to offer the information he sets a price p. Player C observes p, decides whether to buy or not, and if he buys the message he takes a corrective action $d_C \in \mathbb{R}$ against agent A. For tractability and realism, our model is constructed so that receiving an empty message, or not purchasing it in the first place, is uninformative. This is detailed below.

3.1.1 Players' payoffs as functions of B's and C's actions

In this subsection, we consider the players' payoffs as they arise from the two economic decisions by players B and C, d_B and d_C, respectively. We write all three players' payoffs as depending on some bliss points, which are

---

[7]The cost reflects the effort that A exerts when describing her information or when attempting to receive B's attention. One might think of A's choice simply as how often A repeats her message (or how fast she sends it), where sending more (faster) messages to B will increase the likelihood that he will register them (on time to react). This interpretation relates naturally to the convexity of cost.

implicitly defined by squared payoff reductions. (The parameters κ, c_A and c_B are further discussed below the formulae.)

$$\pi_A = -(1 - \kappa)(d_B - \theta_A)^2 - \kappa d_C^2 - c_A(x_A) \tag{1}$$

$$\pi_B = -(d_B - \theta_A)^2 + p - c_B \tag{2}$$

$$\pi_C = -(d_C - \theta_C \theta_A)^2 - p \tag{3}$$

Player A has different bliss points for B's and C' actions. She wants B's action d_B to match her private information θ_A and ideally has C choosing $d_C = 0$. This discrepancy generates her main tradeoff and the parameter $\kappa \in [0, 1]$ governs the weights of her two objectives.[8] Players B and C do not care about each other's actions but each have a bliss point only for their own action. B's bliss point for his decision is the same as A's bliss point, θ_A. It is A's private information and we refer to it as A's "preference type". We assume that θ_A's distribution is strictly increasing, differentiable, and symmetric around zero: $\theta_A \sim F_{\theta_A}$ with existing $f_{\theta_A}(\theta_A) \equiv \frac{dF_{\theta_A}}{d\theta_A}(\theta_A)$ and with $F_{\theta_A}(\theta_A) = 1 - F_{\theta_A}(-\theta_A)$ for all θ_A in the support[9] which is assumed to be an interval $[-\bar{\theta}_A, \bar{\theta}_A]$. To make the analysis interesting we assume that $\bar{\theta}_A$ is large enough such that at least some types would provide strictly positive amounts of information in the absence of leaks: $\bar{\theta}_A^2(1 - \kappa) > c_A'(0)$.

Player C's bliss point for his own action is unknown to both players A and B. It is a linear function of A's preference type, $\theta_C \theta_A$, where θ_C is C's private information. That is, C knows how much he cares about θ_A but he needs knowledge of θ_A in order to know how to set d_C optimally. We assume that θ_C's distribution F_{θ_C} is strictly increasing and differentiable everywhere on its support, which is an interval $[0, \bar{\theta}_C]$. We leave F_{θ_C} unrestricted except that we impose regularity on the square of θ_C, θ_C^2, by assuming that its hazard rate is strictly increasing.[10]

---

[8]Letting all bliss points be real-numbered is akin to assuming that the action spaces of B and C are identical. We make this simplifying assumption although it is unlikely to be literally true in many real-world examples. We point out that different action spaces can often be projected onto a one-dimensional scale, which may e.g. measure the preferred degree of job focus.

[9]Notice that A's bliss point for C's action being equal to zero can be assumed without loss of generality: for any other commonly known bliss point $b_{A,C} \in \mathbb{R}$ we can normalize $b_{A,C}$ to be zero if the distribution of $\theta_A - b_{A,C}$ is symmetric around 0.

[10]The assumption is used as a sufficient condition for uniqueness of the monopoly price.

The above assumptions imply that A would like to tell B about θ_A but prefers C to take actions close to zero—which is the situation we had set out to model. Notice that due to the symmetry of F_{θ_A}, the ex-ante optimal action of C is zero—in the absence of obtaining information C will not (and perhaps cannot) take a punitive action. Finally, notice that, as already discussed above, the model is restricted to situations where it is common knowledge that C is not on A's side. While this is restrictive, it is often the relevant case for issues of privacy: θ_A is A's secret vis-a-vis C.[11]

3.1.2   Information transmission

Both A and B choose whether or not to convey information about θ_A. A chooses whether or not to send B a stochastic message $m_A \in \{\theta_A, \emptyset\}$ that is equal to θ_A with probability x_A and is empty with probability $1 - x_A$. A's only choice is the precision x_A of her message. She incurs a cost $c_A(x_A)$ that is increasing, twice continuously differentiable and convex and satisfies $c_A(0) = 0$ and $c'_A(0) > 0$.

Player B observes the message m_A and, apart from making his choice of d_B, he is also a potential information trader: B decides to offer relaying m_A to C or not, at cost $c_B \in \mathbb{R}^{++}$. We denote this binary decision by $s_B \in \{0, 1\}$. If he offers m_A, i.e. $s_B = 1$, he incurs cost c_B and commits to providing m_A in return for price p. c_B reflects the cost of obtaining and supplying data in the market. It is assumed to be irreversible, such that B incurs the cost even if C rejects the offer. We also assume that p and c_B enter the payoffs additively, as described in (2) and (3).

3.1.3   Naivete

As the last ingredient for our model we introduce two different cognitive types of player A, denoted by τ_A, with possible values n ("naive") and s ("sophisticated"). The naive type occurs with probability α and believes that the cost c_B is higher than it actually is, $\hat{c}_{B,n} > c_B$. As described in the introduction, information flows from B to C may thus appear to stem

---

We argue that it is equally natural to make regularity assumptions on θ_C^2 or on θ_C. For a direct interpretation of θ_C^2, note that is the factor by which θ_A^2 needs to be multiplied to arrive at the utility difference that C imposes on A by choosing his optimal d_C.

[11]This assumption can also be generated by a more symmetric model where C may be either on A's side or not, but has a verifiable technology to prove to A if he is on her side.

from a "leak". For example, a relevant extreme case is $\hat{c}_{B,n} = \infty$, equivalent to assuming that cognitive type n of Player A ignores the possibility that a market for information exists at all. But overestimation of cost may also appear to a milder extent.

We model this dichotomy of belief types in a game with non-common priors. Although only one cost level is true (and known by players B and C), both cost levels of player B are relevant for all players because each cognitive type of player A believes in one of them. We therefore define two possible cost types of player B, $\tau_B \in \{n, s\}$ with cost levels $c_{B,\tau_B} \in \{\hat{c}_{B,n}, \hat{c}_{B,s}\}$, and introduce simple (degenerate) non-common prior beliefs where both the sophisticated type of player A as well as players B and C expect $\hat{c}_{B,s} = c_B$ with probability 1 while the naive type of player A expects $\hat{c}_{B,n}$ with probability 1. These beliefs are common knowledge, that is, all higher-order beliefs are "correct": the different player types agree to disagree in that cognitive type $\tau_A = n$ thinks of $\tau_A = s$ as being paranoid, whereas cognitive type $\tau_A = s$ and players B and C regard type $\tau_A = n$ as being naive and careless.[12] Our welfare analysis will always be based on the true cost, $\hat{c}_{B,s} = c_B$.

Player A's full type profile is (θ_A, τ_A) and we assume that A's preference type θ_A is independent of A's cognitive type τ_A. We also assume that there exist no other stochastic dependence of types between the three players.

3.2 Monopoly

In this subsection, we solve the model under the assumption that there is only one agent who has access to A's message and can sell it to the market, i.e., there is only one player B. Subsection 4.3 discusses the case of oligopoly. Under monopoly, the sequence of moves in the game is as follows.

$t = 0$: All players learn their types.

$t = 1$: A of type (θ_A, τ_A) chooses $x_{A,\tau_A}(\theta_A) \in [0, 1]$.

---

[12]A different modelling approach would have the naive types of player A not understand that other players have different beliefs. For example, naive As might believe that other players believe that all player As are like her. In our model, a naive player A is naive about the costs of information transmission but is sophisticated otherwise. She is not "unaware" like the naive types in the literature on shrouded attributes (Gabaix and Laibson, 2006, Heidhues et al., 2012).

$t = 2$: Nature chooses the content of m_A, with probabilities $x_{A,\tau_A}(\theta_A)$ and $1 - x_{A,\tau_A}(\theta_A)$, and B (of type τ_B) observes it. B chooses d_B and $q_{B,\tau_B} \in [0,1]$, the probability of offering message m_A for sale. If B offers m_A, he quotes price p and incurs cost $\hat{c}_{B,\tau_B}$.[13]

$t = 3$: If p was quoted by B in $t = 2$, C observes it and chooses $buy_C(\theta_C) \in \{0, 1\}$, indicating whether or not to buy m_A. If he buys m_A, he observes it and then chooses $d_C(\theta_C)$. All payoffs are realized.

We restrict attention to the game's symmetric Perfect Bayesian Nash equilibria, where "symmetric" refers to the preference type θ_A: we consider equilibria with $x_{A,\tau_A}(\theta_A) = x_{A,\tau_A}(-\theta_A)$ for all (θ_A, τ_A). This restriction to symmetric equilibria implies that neither B nor C adjust d_B and d_C upon observing $m = \emptyset$. (Hereafter, we simply refer to equilibria, dropping "symmetric Perfect Bayesian Nash".)

An important property of the equilibrium set stems from the observation that there are no equilibria in which B's profit from selling to C depends on the value of m_A. This is because C cannot verify B's information before making his purchasing decision about m_A: if there were equilibria where B's profit was different for different values of m_A then B could always pretend vis-a-vis C to have the information that induces the highest equilibrium payoff. This property implies that in equilibrium C does not adjust his beliefs upon observing p and chooses $d_C = 0$ in case he does not buy the message or when it turns out to be empty.

With these observations on the equilibrium set, we can state our first proposition, describing the behavior of C in a monopoly market. All proofs are in the appendix.

Proposition 1: Equilibrium behavior of player C There exists a unique $\theta_C^* > 0$ such that if m_A is offered in symmetric equilibrium then player C buys it if and only if $\theta_C \in [\theta_C^*, \bar{\theta}_C]$.

The proposition is important for our analysis for two reasons, one substantive, one technical. On a substantive level, it shows that A has to brace herself for the worst. The market allocates the information to those types

---

[13]Strictly speaking, we should also index d_B and p with B's type τ_B. But in equilibrium both types of B choose identical values for these choice variables, hence we drop the indices. As another preview that is relevant for notation, A's and C's equilibrium strategy will not involve randomization for almost all types. We therefore introduce mixed-strategy notation only for player B.

of C whose decision value of the information is highest, i.e. those with the largest equilibrium payoff difference between buying and not buying the information. As the incentives of A and C are opposite, the information will be purchased by those types of C who feel most antagonistic towards A.

The technical relevance lies in the fact that the critical θ_C^* is independent of other aspects of the equilibrium that is played, i.e. θ_C^* is constant even across candidate equilibria of different natures. The property is useful for our further analysis because it fixes the expected volume of trade (and the punishment that A will receive from C) conditional on trade occurring. It will translate into the property that the game's equilibrium is unique.

The proof of Proposition 1 relies on the observation that C's willingness to pay is given by the function

$$WTP_C(\theta_C, \Gamma_A) = \theta_C^2 \Gamma_A, \tag{4}$$

where the factor Γ_A describes the equilibrium value of information and depends only on the collection $\{x_{A,\tau_A}(\theta_A)\}_{\theta_A,\tau_A}$:

$$\Gamma_A \equiv \alpha \int_{\theta_A} x_{A,n}(\theta_A)\theta_A^2 dF_{\theta_A} + (1-\alpha) \int_{\theta_A} x_{A,s}(\theta_A)\theta_A^2 dF_{\theta_A} \tag{5}$$

The proof shows that with this form of WTP_C, different candidate equilibrium collections $\{x_{A,\tau_A}(\theta_A)\}_{\theta_A,\tau_A}$ merely scale up and down the demand for information that B faces, leaving the critical buyer type unaffected.

Describing the remainder of C's equilibrium behavior is straightforward: his acquisition induces the choice of his corrective action, d_C, which equals $\theta_C\theta_A$ if and only if C has acquired m_A and it is non-empty, and equals zero otherwise. For simplicity, we will later notate C's equilibrium behavior by the phrase "C follows Proposition 1" but mean to imply that his choice d_C is also optimal in the described way.

Having solved the last part of the game we consider the truncated game played by A and B. This truncated game has the flavor of a simple inspection game. One the one hand, if A expects that B is likely to sell her message, she will be cautious: she withholds information and thereby reduces the market value of information Γ_A to the point where B's costs are no longer covered such that the market for information may break down. On the other hand, if A expects that there is no market for her information, she will be careless and send her information with maximum precision, raising Γ_A and creating an incentive for C to purchase at prices that make provision worthwhile

13

for B. The equilibrium balances these incentives and, for many parameter constellations, involves mixed strategies. More precisely, the equilibrium predicts the use of mixed strategies for B and a collection of (pure strategies) $\{x_{A,\tau_A}(\theta_A)\}_{\theta_A,\tau_A}$ that make B indifferent between selling and not selling.

Technically, we proceed as follows. As θ_C^* is constant, each type of player A only requires a belief about the probability with which B sells her message, in order to determine her optimal precision. Knowing that B optimally chooses $d_B = \theta_A$ whenever he observes a nonempty message, and $d_B = 0$ otherwise, A' problem is to maximize

$$E[\pi_{A,\tau_A}(\theta_A)] = -\theta_A^2(1-\kappa)(1-x_{A,\tau_A})-\theta_A^2\kappa\hat{q}_{B,\tau_A}x_{A,\tau_A}\int_{\theta_C^*}^{\bar{\theta}_C}\theta_C^2 dF_C(\theta_C)-c_A(x_{A,\tau_A})$$

(6)

with respect to x_{A,τ_A}, where $\hat{q}_{B,\tau_A}$ is type τ_A's belief about q_B. Fixing a pair $(\hat{q}_{B,n}, \hat{q}_{B,s})$ thus induces the collection $\{x_{A,\tau_A}(\theta_A)\}_{\theta_A,\tau_A}$ of best replies to the pair. Conversely, any such collection pins down what the two B types can earn if they sell the information at their given cost and, hence, it pins down their optimal decisions $(q_{B,n}, q_{B,s})$. An equilibrium solves this fixed-point problem from $(\hat{q}_{B,n}, \hat{q}_{B,s})$ to $(q_{B,n}, q_{B,s})$. To find it, we consider *B's equilibrium revenue conditional on selling to C*: we denote by $H(q_{B,n}, q_{B,s})$ the expected monopolistic revenue that B would receive from selling m_A to C, given that simultaneously (i) A believes, and C knows that A believes, that the two types of B offer m_A for sale with probabilities $(\hat{q}_{B,n}, \hat{q}_{B,s}) = (q_{B,n}, q_{B,s})$, (ii) C's behavior follows Proposition 1, and (iii) A maximizes (6). It is given by

$$H(q_{B,n}, q_{B,s}) = WTP_C(\theta_C^*, \Gamma_A(q_{B,n}, q_{B,s}))(1 - F_C(\theta_C^*)),$$

where the notation $\Gamma_A(q_{B,n}, q_{B,s})$ indicates that the market value of information depends on $(q_{B,n}, q_{B,s})$.

Notice that $H(q_{B,n}, q_{B,s})$ is identical for both types of B as their costs do not affect their price setting. Notice also that $H(q_{B,n}, q_{B,s})$ does not depend on whether or not B actually offers m_A for sale. In other words, it is well-defined only from A's and C's behavior. A simple but important implication of (6) is that all types of A will send (weakly) more information as $(q_{B,n}, q_{B,s})$ decreases, which allows establishing that H is decreasing in its arguments. The appendix proves this property and shows how it results in Proposition 2.

Proposition 2: Equilibrium behavior of players A and B There exists a unique symmetric equilibrium. For no constellation of parameters does

the market for information shut player A up: A sends nonempty m_A with strictly positive probability. A maximizes (6) with an equilibrium belief about B's strategy, $(\hat{q}_{B,n}, \hat{q}_{B,s}) = (q_{B,n}, q_{B,s})$ that is characterized by the following, where $0 \leq H(1,1) < H(0,1) < H(0,0)$:

I If $\hat{c}_{B,n} \leqslant H(1,1)$, then $q_{B,n} = q_{B,s} = 1$.

II If $H(1,1) < \hat{c}_{B,n} < H(0,1)$, then $q_{B,n} \in (0,1)$ and $q_{B,s} = 1$.

III If $\hat{c}_{B,s} \leqslant H(0,1) \leqslant \hat{c}_{B,n}$, then $q_{B,n} = 0$ and $q_{B,s} = 1$.

IV If $H(0,1) < \hat{c}_{B,s} < H(0,0)$, then $q_{B,n} = 0$ and $q_{B,s} \in (0,1)$.

V If $H(0,0) \leqslant \hat{c}_{B,s}$, then $q_{B,n} = q_{B,s} = 0$.

Figure 1 illustrates the five regions of cost parameters that lead to the different equilibrium behaviors of B. Notice that differently from Figure 1's illustration, it may be that $H(1,1) = 0$, in which case region I fails to exist. This possibility is further discussed below.

We now describe the five regions, one by one. The description also uses some auxiliary results from the appendix: Proposition A.1 shows that a naive A provides more information than a sophisticated A, i.e. $x_{A,n}(\theta_A) \geq x_{A,s}(\theta_A)$ for all θ_A, and that more "interesting" types of A—those with θ_A further away from zero—provide more information, i.e. $x_{A,\tau_A}(\theta_A)$ increases weakly in $|\theta_A|$ for all τ_A. Proposition A.2 describes how the players' equilibrium payoffs vary with the cost parameters. For the equilibrium description it is also useful to denote by $\underline{x}_A(\theta_A, \tau_A)$ and $\overline{x}_A(\theta_A, \tau_A)$, respectively, the *minimal and maximal information provision* that player A of type (θ_A, τ_A) chooses in response to the most pessimistic and optimistic expectations about B's probability of selling to C, $\hat{q}_{B,\tau_A} = 1$ and $\hat{q}_{B,\tau_A} = 0$. The value of $\underline{x}_A(\theta_A, \tau_A)$ may or may not be zero, just as $\overline{x}_A(\theta_A, \tau_A)$ may or may not be equal to one, depending on the model's parameters.

I (if it exists): For cost parameters in this region, both types of A anticipate that B sells m_A for sure. All types of A therefore provide the minimal precision $\underline{x}_A(\theta_A, \tau_A)$. The proof of Proposition 2 shows that region I exists iff there exist some types of A whose minimal information provision $\underline{x}_A(\theta_A, \tau_A)$ is strictly positive, and that this happens iff it is true for the most extreme preference type, $\overline{\theta}_A$. More precisely, the region exists if

$$1 - \kappa\left(1 + \int_{\theta_C^*}^{\overline{\theta}_C} \theta_C^2 dF_C(\theta_C)\right) \geq \frac{c'(0)}{\overline{\theta}_A^2} \tag{7}$$

15

$\hat{c}_{B,n}$

V

IV

$q_{B,n} = 0$
$q_{B,s} = 0$

$q_{B,n} = 0$
$0 < q_{B,s} < 1$

III

$q_{B,n} = 0$
$q_{B,s} = 1$

II
$0 < q_{B,n} < 1$
$q_{B,s} = 1$

I

$q_{B,n} = 1$
$q_{B,s} = 1$

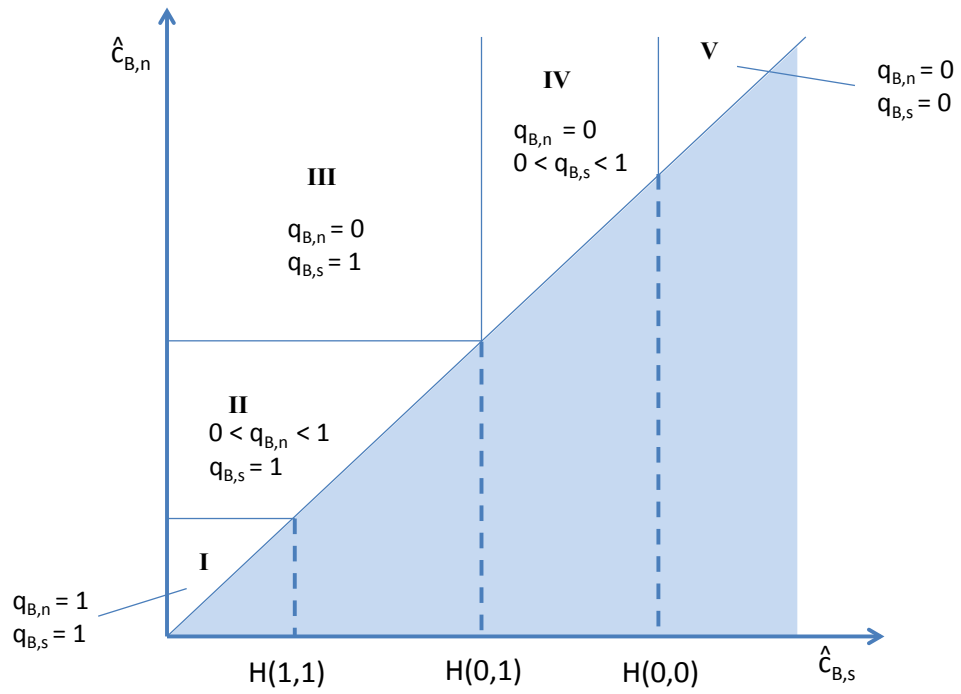H(1,1)   H(0,1)   H(0,0) $\hat{c}_{B,s}$

Figure 1: Equilibrium behavior of player B

and only regions II to V exist otherwise.

II: In this parameter region, the sophisticated A continues to correctly anticipate that B sells for sure and she hence chooses $\underline{x}_A(\theta_A, s)$. But the naive A is more optimistic (and less careful) about the behavior of his imaginary opponent, high-cost B. The condition $H(1,1) < \hat{c}_{B,n}$ implies that $\hat{q}_{B,n} < 1$ and every θ_A-type of the naive A plays her (pure-strategy) best response to $\hat{q}_{B,n}$, which lies strictly above $\underline{x}_A(\theta_A, n)$ for some θ_A.[14] The high-cost B's cost, $\hat{c}_{B,n}$, is larger than in region I and he therefore has an incentive to reduce $q_{B,n}$ as long as $H(q_{B,n}, 1) < \hat{c}_{B,n}$. In equilibrium, he makes himself indifferent by setting $q_{B,n}$ such that $H(q_{B,n}, 1) = \hat{c}_{B,n}$. The (true) low-cost B's profits in region II exceed those of region I as the additional information that the naive A now creates additional demand by C.

III: In this region the naive type of player A is even more optimistic about the cost of information transmission. She feels uninhibited in providing information, $x_{A,n}(\theta_A) = \overline{x}_A(\theta_A, n)$. The sophisticated A anticipates that B sells the information for sure and chooses the minimal information provision $\underline{x}_A(\theta_A, s)$. Player B sells the minimal information provided by sophisticated As, plus the maximal information provided by the naive As. In terms of profits this is B's best parameter region.

IV: While naive A still chooses maximal $\overline{x}_A(\theta_A, n)$ (as she assume that B will never sell), now the sophisticated A plays a game of cat and mouse with low-cost B. The logic for the interaction between the sophisticated A and (real) low-cost B is analogous to the mixed-equilibrium logic of region II: sophisticated A provides just enough information to make B indifferent between selling and not, and B sells with low enough probability keep A willing to provide *some* information. But there is an important difference regarding the actual profits of (low-cost) B. In the game of cat and mouse in region IV, the sophisticated A correctly considers the revenue that B can generate from selling information sent by the naive A. Consequently, in expectation, the low-cost B makes zero profits in this equilibrium.

V: While desirable from the perspective of As, this is the least interesting equilibrium region. B's true costs are prohibitive and so are the even higher costs imagined by the naive A player. There is no market for information. All types of A choose maximal information precision $\overline{x}_A(\theta_A, n)$ and achieve their highest possible payoff, while player B earns nothing as a trader of

---

[14](6) implies that only a zero-measure set of naive A types are indifferent between different precision levels.

information.

We now consider a selection of comparative statics results. As indicated above, the appendix contains a comprehensive result on comparative statics, Proposition A.2, regarding the players' payoffs when moving within and between the different regions. In reality, such moves may be generated by regulatory intervention that affect the actual cost of B entering the market for information, $\hat{c}_{B,s} = c_B$, and/or the perceived cost $\hat{c}_{B,n}$. The following corollary considers the consequences of increasing both $\hat{c}_{B,s}$ and $\hat{c}_{B,n}$ by constant proportions, that is, along some ray $\hat{c}_{B,n} = \lambda \hat{c}_{B,s}$ with $\lambda > 1$. In Figure 1, this corresponds to moving northeast on a straight line that starts at the origin and lies above the 45-degree line.[15]

Corollary 1 For sufficiently low $(\hat{c}_{B,n}, \hat{c}_{B,s})$,

(i) a modest proportional increase in $(\hat{c}_{B,n}, \hat{c}_{B,s})$ weakly reduces naive A's payoff, for all θ_A, without increasing sophisticated A's payoff;

(ii) a sufficiently large proportional increase in $(\hat{c}_{B,n}, \hat{c}_{B,s})$ weakly increases the payoffs of all types of player A.

Clause (i) demonstrates that a half-baked regulatory invention would only harm the naive player A, as she would feel safe too soon. The result derives from the observation that a modest proportional increase in $(\hat{c}_{B,n}, \hat{c}_{B,s})$ corresponds to a northeast move within region II or a move from region I to region II. The cost level weakly increases the potential equilibrium revenue H because high-cost B has to be indifferent in region II. By weak monotonicity of H, this means that $q_{B,n}$ is reduced, inducing a weakly higher $x_{A,n}$. Since the true probability of selling is $q_{B,s} = 1$, this leads to a payoff reduction (otherwise, sophisticated A would increase $x_{A,s}$, too). Clause (ii) refers to a strong measure of privacy protection, moving from region I or II to region V, where all types of A have maximal payoffs.

Somewhat counter-intuitively, it may *not* be possible to induce Pareto improvement through an increase in $(\hat{c}_{B,n}, \hat{c}_{B,s})$, relative to a situation where A's personal data is relatively unprotected. One may have thought that stronger data protection enables the players to move from a completely silent equilibrium to one where an intermediate level of information is shared and traded,

---

[15]The result is presented as a corollary as it is derived from Proposition 2 and Proposition A.2 in the appendix.

to everyone's benefit. But as Proposition 2 shows, the equilibrium never allows for complete silence: in equilibrium regions I and II, players B and C benefit from the naive players who chatter away. Moving towards regions IV and V would therefore not necessarily improve B's and C' payoffs.[16]

Our next corollary concerns variations in the share of naive players, which might be affected by policy interventions such as educational campaigns about the actual costs of surveillance and data provision.

Corollary 2 A reduction in α

(i) increases information provision by any given type of player A:

$$\frac{dx_{A,n}(\theta_A)}{d\alpha} \leq 0 \text{ and } \frac{dx_{A,s}(\theta_A)}{d\alpha} \leq 0, \forall \theta_A$$

(ii) benefits the sophisticated A: $\frac{d\pi_{A,s}(\theta_A)}{d\alpha} \leq 0, \forall \theta_A$

(iii) can harm the naive A: $\exists \alpha', \alpha$ such that $\alpha' > \alpha$ and $\pi_{A,n}(\theta_A)$ is strictly larger at α' than at α, for all θ_A.

The corollary shows how all types of A compensate for a large prevalence of naivete. A larger share of naive players wets the appetite of C to buy the available information, and hence the likelihood that B is willing to sell. As a result, both naive and sophisticated A reduce their information provision (clause (i)). The overall value of the available information Γ_A, however, increases in response to a larger α because naive agents provide more information in the aggregate. (Otherwise, if Γ_A was reduced, then all types of A would increase their x_A. Since naive A provides more information than sophisticated A, and naive A's share increases, the resulting aggregate level of Γ_A would have to increase, a contradiction.) For the sophisticated A, this effect translates into a reduction of the equilibrium payoff, in response to larger α. In contrast, the increase in α may *help* the naive players. This is possible because an increase in α affects the boundaries of the equilibrium regions and the increase in α may thus corresponds to a change from the equilibrium prediction in region I to the equilibrium prediction in region II:

---

[16]This does not mean that no Pareto improvement is possible through manipulations of $(\hat{c}_{B,n}, \hat{c}_{B,s})$, but we believe that one would need stronger assumptions for it.

the increase in α makes naive player A behave more cautiously and reduce their loss.

Our final, and brief, comparative statics exercise examines a variation in the degree to which player C cares about A's information. What if C becomes more or less aggressive towards A? That is, we consider the consequences of scaling the distribution of θ_C. Let $\widetilde{\theta}_C = \mu\theta_C$ and $\widetilde{F}(\widetilde{\theta}_C) = F(\theta_C)$. Player C's willingness to pay for the information now becomes $WTP(\widetilde{\theta}_C) = \mu^2\theta_C^2\Gamma_A$. Scaling the distribution of C types is therefore equivalent to shifting Γ_A and leaves the selection of buyers in B's monopoly untouched: for the distribution $\widetilde{F}$ the marginal type to which B sells is simply given by $\widetilde{\theta}_C^* = \mu\theta_C^*$. The consequences for A's choice of information precision are, hence, straightforward. For $\mu > 1$ ($\mu < 1$) her information becomes more (less) valuable in the market and the potential punishment more (less) severe. Thus, she will weakly reduce (increase) her precision and will be weakly worse (better) off.

4 Extensions and robustness

4.1 Naivete about the market

An alternative and perhaps equally plausible form of naivete by player A concerns the misappreciation of the market mechanism that allocates the information to player C. Naive As may not grasp that, conditional on information being passed on, it will reach the most aggressive C-types (those above θ_C^*) and not just a random selection of C-types. Modeling this alternative is relatively straightforward, applying the common logic of Jehiel's (2005) analogy-based expectations equilibrium and Eyster and Rabin's (2005) cursed equilibrium. Differing from Section 3, we here assume that naive A has a correct perception of B's cost c_B (just as sophisticated A) and attaches the correct belief R to the event that their message is intercepted and bought by player C. That is, we impose the equilibrium condition

$$R = q_B \int_{\theta_C^*}^{\overline{\theta}_C} dF_C.$$

such that the naive A has rational expectations about the likelihood of her message to be acquired by C. (Notice that in this variation all B types know the the true costs such that we do not require an additional index for q_B.)

However, when naive A thinks about the precision with which to send her information she now maximizes

$$-\theta_A^2(1 - \kappa)(1 - x_{A,n}) - \theta_A^2 \kappa x_{A,n} R \int_0^{\bar{\theta}_C} \theta_C^2 dF_C - c_A(x_{A,n}).$$

While her beliefs about the likelihood of being punished by C are correct, she systematically underestimates the size of the expected punishment because she does not understand that the market allocates the information systematically to high types. Hence, as in the main model, naive A will send more information than sophisticated A.

The analysis of this variation proceeds in a similar way as the analysis of Section 3. Crucially, Proposition 1 still holds for any possible belief R. That is, independent of the nature of the equilibrium and independent of R, there is a critical θ_C^* defining a lower bound for which C types acquire the information. (Only the value of Γ_A differs from that in Section 3, and depends on R, as naive A solves a different optimization problem.) Consequently, in equilibrium R depends linearly on q_B and we denote this equilibrium mapping as $R(q_B)$.

The remaining analysis is even simpler than in Section 3, as we now have only one cognitive type of A and one type of B to consider (the true, low-cost B). For given q_B, each θ_A-type of the sophisticated player A maximizes (6). The probability q_B thus determines the equilibrium revenue from selling m_A to C. We denote it as $H(q_B)$ and point out that H depends on q_B through two channels, as a direct effect on the equilibrium choice of the sophisticated A and through the equilibrium effect on the choice of the naive A. Similar to Section 3, one can now show that H decreases in q_B, which ensures uniqueness. The equilibrium strategies again depend on the relative size of revenue H versus cost c_B, with three regions similar to regions I, III and V of Proposition 2. For $c_B > H(0)$, the equilibrium region applies that player A likes best where $q_B = 0$ and all As choose maximal information precision (analogous to region V). Similarly, a region analogous to region I may or may not exist for $c_B < H(1)$, where B's strategy prescribes $q_B = 1$ and A chooses minimal information precision.[17] Finally, for intermediate levels of

---

[17]The region exists if

$$1 - \kappa(1 + \int_0^{\bar{\theta}_C} \theta_C^2 dF_C(\theta_C)) \geq \frac{c'(0)}{\bar{\theta}_A^2}$$

21

B's costs, $H(1) \leq c_B \leq H(0)$, we obtain a mixed-strategy equilibrium where the sophisticated types of player A make player B indifferent between selling their messages and not.

In terms of welfare, the picture is also similar to the model of Section 3. Crucially, nothing changes for the welfare of sophisticated A types. Their utility moves hand in hand with their equilibrium information precision and is, hence, increasing in c_B. As before, the sophisticated A's welfare is decreasing in the number of their naive counterparts. As the latter send more information, the former have to reduce their own information precision to keep the Bs indifferent for intermediate c_B.

For naive A the story is, however, slightly more complicated. As long as $q_B > 0$ she provides too much information. As c_B increases (in the interior region), q_B decreases and this has two effects on naive As' welfare. On the one hand, she benefits from the fact that the net marginal value of sending information increases. On the other, she incurs an additional loss through providing an even higher uncalled for information precision. Either of these two effects can dominate depending on the curvature of $c_A(x_A)$.

4.2 Commitment from B

In the main model we assume that B, in his role as seller, can as easily adapt as all other players. However, in many situations B, as the interceptor and seller of information, may be forced to invest in a technology for dealing with data which would be tantamount to assuming that B has to move first committing to some q_B. The consequences of such commitment are that B is more cautious in his dealings with A, being careful not to reduce A's willingness to provide information. A credible commitment to protect data security may thus be is an improvement for both A and B. (And C, too, may benefit from more information being sent.) Moreover, B, as interceptor and trader of information, has an incentive to make his doing public.

However, it may also not be in B's interest to address data security issues too much, in order not to alert naive A about his own incentives to leak information. We do not address the latter question in any formal way but here only briefly consider the equilibrium effect of commitment.

The simplest way of illustrating the effects of such commitment is by

---

i.e., if the naive A provides strictly positive information precision for her most pessimistic belief, $R = 1$.

focussing on a setup where there are no naive types of player A. The sequence of events with commitment is as follows. First, B chooses q_B, then A chooses x_A and, finally, A's message is traded. Nothing fundamental changes for the last stage, relative to section 3, and a variant of Proposition 1 still holds. However, now player B maximizes

$$q_B[H(q_b) - c_B]$$

where $H(q_B)$ is again the expected revenue can reap by selling m_A. The first-order condition is

$$H(q_B) - c_B + H'(q_B)q_B = 0$$

with $H'(q_B) < 0$. B therefore chooses q_B in a way that ensures positive profits, that is, we will have $H(q_B) > c_B$. For intermediate values of c_B (that is, for the most interesting case), this is beneficial for B compared to the model without commitment. It is also beneficial for A because $H' < 0$ ensures that B's q_B with commitment will be strictly smaller than without. Consequently, A sends more information in a model with commitment than without and will be strictly better off (for intermediate levels of c_B).

4.3   Competition

What happens when there are multiple interceptors/sellers of information? We model seller competition by assuming that simultaneous to player A's choice of precision x_A, N identical firms $B_1, ..., B_N$ make entry decisions. If firm i enters the market, she pays c_B as irreversible entry cost and observes m_A. Then all firms observe each others' entry decisions and subsequently the entering firms simultaneously choose a price at which C can buy m_A. We assume that C buys from at most one firm B_i, inducing Bertrand competition among the entering firms. Otherwise, the model is identical to that of Section 3.

   In case of multiple entries, Bertrand competition induces a market price of zero. If exactly one firm enters, it can set a monopoly price, analogous to Section 3. The resulting entry game will have multiple asymmetric equilibria but we focus on the unique symmetric mixed-strategy equilibrium, where each firm enters with some non-zero probability. Anticipating the mixed-strategy equilibrium, player A simply has to consider the likelihood of two possible cases when determining her precision x_A—the likelihood of a single

entry occurring, in which case her information will be sold to players C with $\theta_C \geqslant \theta_C^*$, and the likelihood of multiple entry, in which case her information will be sold to *all* types of player C. Analogous to Section 3, naive and sophisticated A can differ in their beliefs about the likelihood of these events. Let A's beliefs about the probability for a single entry be denoted by $s_{\tau_A}^1$ and her belief about multiple entries by $s_{\tau_A}^2$. Player A then maximizes

$$-\theta_A^2(1-\kappa)(1-x_{A,\tau_A})-\theta_A^2\kappa x_{A,\tau_A}\left(s_{\tau_A}^1\int_{\theta_C^*}^{\bar{\theta}_C}\theta_C^2 dF_C + s_{\tau_A}^2\int_0^{\bar{\theta}_C}\theta_C^2 dF_C\right)-c_A(x_{A,n})$$

where in equilibrium player A's beliefs, s^1 and s^2, are determined by

$$s_{\tau_A}^1 = Nq_{B,\tau_A}(1-q_{B,\tau_A})^{N-1} \tag{8}$$

and

$$s_{\tau_A}^2 = 1 - s_{\tau_A}^1 - (1-q_{B,\tau_A})^N. \tag{9}$$

For $N = 1$ this contains the earlier monopoly case where $s_{\tau_A}^1 = q_{B,\tau_A}$ and $s_{\tau_A}^2 = 0$. However, now a seller B's probability for selling is determined by the mixed equilibrium of the entry (sub)game. In order to be indifferent between entering and not entering, we need that

$$\hat{c}_{B,\tau_B} = (1-q_{B,\tau_B})^{N-1}\mathcal{H}(q_{B,n}, q_{B,s}) \tag{10}$$

holds, where $\mathcal{H}(q_{B,n}, q_{B,s})$ denotes the monopoly revenue a seller achieves in equilibrium when he is the only entrant.[18]

Similar to Section 3's arguments about monotonicity of H, it is now useful to show that total expected punishment $s_{\tau_A}^1\int_{\theta_C^*}^{\bar{\theta}_C}\theta_C^2 dF_C + s_{\tau_A}^2\int_0^{\bar{\theta}_C}\theta_C^2 dF_C$ is monotone in q_{B,τ_A}.[19] Hence, information provision decreases monotonically in q_{B,τ_A} and so do the value of information, Γ_A, and revenue $\mathcal{H}$. Consequently, the right-hand side of equation (10) also decreases in q_{B,τ_A} ensuring uniqueness of the symmetric equilibrium in the entry (sub)game, for each of the two possible types of firms.

---

[18]Notice that we use calligraphic $\mathcal{H}$ instead of the earlier H to signify that the revenue that arises in such an endogenous monopoly is different from the revenue achieved in the earlier exogenous monopoly case, simply because player A's maximization problem is now different: she faces monopoly as only one possible outcome of the entry decisions.

[19]Simply use the fact that the punishment is more severe if Bertrand competition ensues.

The resulting equilibrium configuration for oligopoly is similar to the monopoly case depicted in Figure 1. However, the analogue to old equilibrium region I (with $\hat{c}_{B,s}, \hat{c}_{B,n} < \mathcal{H}(1,1)$) no longer exists because the right hand-side of (10) can be made arbitrarily small through its first term. That is, an equilibrium with certain entry by all firms does not exist (it would generate negative profits for all firms). For $\hat{c}_{B,s}, \hat{c}_{B,n} < \mathcal{H}(0,1)$ we now obtain a region where both the (real) low-cost sellers and the (imagined) high-cost sellers mix. For $\hat{c}_{B,n} \geq \mathcal{H}(0,1)$ and $\hat{c}_{B,s} < \mathcal{H}(0,0)$ we obtain the analogue to the old region IV where low-cost sellers mix and the imagined high-cost sellers abstain from entry. Finally, for $\hat{c}_{B,s}, \hat{c}_{B,n} \geq \mathcal{H}(0,0)$ costs are prohibitive and no entry occurs (the equivalent to the old region V).

Given the similar construction of equilibria it is not surprising that comparative statics with respect to the sellers' costs and the share of naive agents A still hold within equilibrium regions. In particular, higher real costs, $\hat{c}_{B,s}$, benefit all players A while higher imagined (low) costs, $\hat{c}_{B,n}$, can harm naive types of A (because they underestimate the probability with which their information will be sold). However, regarding our earlier exercise that examined the welfare of A players moving along a ray, $\hat{c}_{B,n} = \lambda \hat{c}_{B,s}$ (with some $\lambda > 1$) differences emerge. The adverse effect of (misguided regulation) moving from the old equilibrium region I to II and III where naive players A get too careless no longer exists. However, within the new region, $\hat{c}_{B,s}, \hat{c}_{B,n} < \mathcal{H}(0,1)$, adverse effects of regulation are still possible depending on the shape of c_A.[20]

While these comparative statics bring nothing much new, we can now examine a new type of comparative statics by varying N, i.e. investigate the effect of more or less competition. Here we find a surprisingly general result: for all parameter constellations, sophisticated A is better off if $N > 1$ (oligopoly) than if $N = 1$ (monopoly). To see this, consider the cost regions in Figure 1, and compare sophisticated A's welfare between monopoly versus oligopoly. In regions I, II and III, monopoly gives sophisticated A the lowest possible payoff, hence oligopoly must be weakly better. In region IV, it must be that in equilibrium $\mathcal{H}(q_{B,n}, q_{B,s})$ exceeds $H(q_{B,n}, q_{B,s})$ (the revenue in the equilibrium of the monopoly game) because otherwise there is no incentive for firms to enter in oligopoly: in the monopoly game, B is indifferent between

---

[20]While increases in costs c_B will reduce the objective probability of information being sold, the wedge between rational information transmission and overprovision of information grows which, if c'_A is flat, can imply that, on balance, naive players A will be worse off.

25

entering and not, i.e. $H(q_{B,n}, q_{B,s}) = \hat{c}_{B,s}$, but to induce entry under the uncertainty of other firms' entry, the possible revenue must be strictly higher than $\hat{c}_{B,s}$. This requires that C's willingness to pay must be strictly higher in the oligopoly game, and thus information provision by A must also be strictly higher. Now observe that naive A provides maximal information in the monopoly game in region IV. Therefore, sophisticated A must be willing to provide more information in the oligopoly game which goes hand in hand with higher payoffs. Finally, in region V, observe that sophisticated A provides maximal information in the monopoly game, and equally so in the oligopoly game: $\mathcal{H}(0,0)$ and $H(0,0)$ coincide, and so does A's maximization problem if the information is never sold in either of the two market structures.

Naive A, in contrast, may be better off under monopoly if $\hat{c}_{B,n} < H(0,1)$, i.e. if regions I or II apply in the monopoly game. This can happen if the mixed-strategy entry game in oligopoly induces naive A to enter with very high probability, relative to the monopoly game. But the opposite can also be true, because under competition naive types may have beliefs that are much better adjusted. To see this point, consider naive As who overestimate costs just a little, $c_{B,n} = c_{B,s} + \delta$. For levels of $c_{B,s}$ that lie slightly below $H(0,1)$, we can have that $c_{B,s} < H(0,1) < c_{B,n}$, in which case the naive A's small aberration has stark consequences: she expects that messages are 100% safe, whereas they are in reality sold with probability 1. In contrast, under Bertrand competition with $N \to \infty$, one can show that the probability of a single entry $s_{\tau_A}^1$, vanishes and the (perceived) probability of multiple entry, $s_{\tau_A}^2$, approaches[21] $1 - \frac{c_{B,\tau_B}}{\mathcal{H}(q_{B,n}, q_{B,s})}$. Therefore, the naive A's belief is much better adjusted under competition with $N \to \infty$: the difference between the information sale's true probability and probability perceived by a naive player is for that case only $\frac{\delta}{\mathcal{H}(q_{B,n}, q_{B,s})}$.

5   Conclusion

Our study highlights the dangers of information leakage through markets. But the analysis also shows that markets for leaks will not necessarily reduce information flows to zero. We illustrate this by returning to the example of mobile phone applications: B may be an app provider, bringing a service to the telephone user but selling his information on to C. The example illus-

---

[21]Simply solve (10) for $q_{B.\tau_B}$ and insert, along with (8) into (9) and take the limit.

trates that app providers need to offer the users a benefit, like a (seemingly) useful software, and that this necessity to engage the customer disciplines the industry. Without receiving a benefit, few rational customers would provide the data.

But how realistic is full rationality in these environments where few people know the intricacies of tracking and technical change is rapid? Presumably, not too many; naive app users might not even notice that they constantly provide information to a hidden data aggregator. As our analysis indicates, it is natural to expect that the industry's larger profits lie in dealing with naive customers who do not withhold information. The data trading industry has strong incentives to create markets where selling information is still unanticipated by originators of information. In markets for new IT products, this is typically the case. But even sophisticated customers suffer because of the externality discussed above: the demand for data interception services increases with the degree of naivete in the population.

In the main analysis, we model the bias in a particular misperception, the over-estimation of cost. But naivete about markets is multi-faceted phenomenon, and one can think of many different models to capture it. One such alternative, where naive consumers do not understand the allocation mechanism that a market provides, is explored in our extensions section. Other equilibrium effects may be equally hard to understand. Take, for example, our finding that player A can be better off if multiple agents eavesdrop on her communication. This result strikes us as rather counterintuitive: if given a choice between having one and multiple agents spying on them, most people might prefer to have just one. This could also make for an interesting experiment. People might underestimate how competition can destroy rents and, thus, renders their communication more safe.

We also examined the scope for regulation to improve outcomes, through measures that affects B's costs of selling A's information. Unsurprisingly, both sophisticated and naive A's welfare can be improved through extreme regulatory measures which render the costs for selling information prohibitive. Less extreme (and, hence, perhaps more relevant) regulatory measures can, however, backfire: they can reduce naive A's welfare without improving the welfare of sophisticated A.

Other regulatory measures might impose constraints on the punishment that C can inflict on A. In oir model, punishments are unbounded but in many of the applications that we discussed, for example, at the workplace, there will be legal limits to what C can do to A. This may be an interesting

avenue for further research.

Finally, a word of caution. We wrote this paper in a way that suggests that the exchange of information between A and B, their joint project, is socially desirable. However, there are, of course, applications that have exactly the same game structure but where society has a collective interest in shutting down the collaboration between A and B because it is in the widest possible sense of a criminal nature. For these applications, all our results are still relevant only that the desirability of outcomes has now the opposite sign. This possibility implies a possible conundrum for regulating information flows on the internet: interference with the market for information is unlikely to have unambiguous positive effects. But this renders further research into the economics of information leaks all the more important.

References

[1] Acquisti, A., and J. Grossklags (2005). *Privacy and rationality in individual decision making.* IEEE Security & Privacy (Jan/Feb): 24-30.

[2] Acquisti, A., C. Taylor and L. Wagman (2014). *The economics of privacy.* Mimeo, Carnegie Mellon University.

[3] Armstrong, M. (2015). *Search and ripoff externalities.* Mimeo, Oxford University.

[4] Armstrong, M., and J. Zhou (2010). *Conditioning prices on search behavior.* Mimeo, University College London.

[5] Arrunada and Vazquez (2006). *When your contract manufacturer becomes your competitor.* Harvard Business Review, 84 (9): 135-145.

[6] Beresford, A.R., D. Kübler and S. Preibusch (2012). *Unwillingness to pay for privacy: A field experiment.* Economics Letters 117: 25-27.

[7] Bergemann, D., and A. Bonatti (2015). *Selling cookies.* American Economic Journal: Microeconomics 7: 259-294.

[8] Börgers, T. (2015). *An Introduction to the Theory of Mechanism Design.* Oxford University Press.

[9] Crawford, V.P., and J. Sobel (1982). *Strategic information transmission.* Econometrica 50: 1431-1451.

[10] Eyster, E., and M. Rabin (2005). *Cursed equilibrium.* Econometrica 73: 16231672.

[11] Gabaix, X., and D. Laibson (2006). *Shrouded attributes, consumer myopia, and information suppression in competitive markets.* Quarterly Journal of Economics 121 (2): 505 540.

[12] Heidhues, P., and B. Koszegi (2014). *Using information about naivete to price discriminate.* Mimeo, ESMT.

[13] Heidhues, P., B. Koszegi and T. Murooka (2012). *The market for deceptive products.* Mimeo, ESMT.

[14] Ho, S. (2009). *Information leakage in innovation outsourcing.* R&D Management 5: 431-443.

[15] Jehiel, P. (2005) *Analogy-based expectation equilibrium.* Journal of Economic Theory 123: 81104.

[16] Jentzsch, N. (2014). *Secondary use of personal data: a welfare analysis.* European Journal of Law and Economics, forthcoming.

[17] Koszegi, B. (2014) *Behavioral Contract Theory.* Journal of Economic Literature 52: 1075-1118.

[18] Norberg, P.A., D.R. Horne and D.A. Horne (2007). *The privacy paradox: Personal information disclosure intentions versus behaviors.* Journal of Consumer Affairs 41 (1): 100-126.

[19] Office of Fair Trading (2013). *The economics of online personalised pricing.* OFT Report 1488.

[20] Schudy, S., and V. Utikal (2014). *Preferences for data privacy: Sharing personal information with close and distant peers.* Mimeo, University of Erlangen-Nuremberg.

[21] Taylor, C. (2004). *Consumer privacy and the market for customer information.* RAND Journal of Economics 35 (4): 631-650.

[22] Tsai, J.Y., S. Egelmann, L. Cranor and A. Acquisti (2011). *The effect of online privacy information on purchasing behavior: An experimental study.* Information Systems Research 22 (2): 254-268.

A   Appendix A: Proofs

Proof of Proposition 1

Consider C's choice of d_C. The uncertainty that he faces stems from the multidimensional randomness in x_A, θ_A and τ_A. For a notation that avoids multiple integrals, we denote by ξ a random variable that governs all of this randomness and we let $\mathfrak{B}_\xi$ be its support and F_ξ its distribution function.

C's optimal strategy d_C^* is a function of (ξ, θ_C). As indicated in the main text, p does not carry any relevant information and the symmetry of $x_{A,\tau_A}(\theta_A)$ around $\theta_A = 0$ implies that C's optimal action is $d_C^* = 0$ whenever he does

not buy m_A or if it is empty. Let $\mathfrak{B}_\xi^0$ be the subset of $\mathfrak{B}_\xi$ where $d_C^* = 0$ if C buys m_A. The "other" values of ξ, $\xi \notin \mathfrak{B}_\xi^0$, are those where C, if he buys m_A, observes a nonempty and nonzero value $m_A = \theta_A \neq 0$, and optimally sets $d_C^*(\theta_C) = \theta_C \theta_A$. We can then write C's expected payoff from buying m_A and responding to it with his optimal strategy d_C^* as

$$E[\pi_C(buy_C = 1, d_C^*, \theta_C)] = \int_{\xi \in \mathfrak{B}_\xi} -(d_C^* - \theta_C \theta_A)^2 dF_\xi - p$$

$$= \int_{\xi \in \mathfrak{B}_\xi^0} -(0 - \theta_C \theta_A)^2 dF_\xi + \int_{\xi \notin \mathfrak{B}_\xi^0} -(\theta_C \theta_A - \theta_C \theta_A)^2 dF_\xi - p$$

$$= \int_{\xi \in \mathfrak{B}_\xi^0} -(\theta_C \theta_A)^2 dF_\xi - p$$

C's payoff from not buying m_A is

$$E[\pi_C(buy_C = 0, d_C^*, \theta_C) = \int_{\xi \in \mathfrak{B}_\xi} -(\theta_C \theta_A)^2 dF_\xi$$

and his willingness to pay for m_A is therefore:

$$WTP_C(\theta_C) = E[\pi_C(buy_C = 1, d_C^*, \theta_C)] + p - E[\pi_C(buy_C = 0, d_C^*, \theta_C)]$$

$$= \int_{\xi \in \mathfrak{B}_\xi^0} -(\theta_C \theta_A)^2 dF_\xi - p + p - \int_{\xi \in \mathfrak{B}_\xi} -(\theta_C \theta_A)^2 dF_\xi$$

$$= \theta_C^2 \int_{\xi \notin \mathfrak{B}_\xi^0} \theta_A^2 dF_\xi$$

Now we give up the ξ-Notation and ask about the random outcomes that correspond to $\xi \notin \mathfrak{B}_\xi^0$. When is m_A nonempty? All of the randomness in x_A, θ_A and τ_A is independent by assumption. Aggregating across τ_A yields the θ_A-specific precision of m_A. Aggregating further across θ_A gives m_A's overall average probability of being nonempty as

$$\int_{\xi \notin \mathfrak{B}_\xi^0} dF_\xi = \alpha \int_{\theta_A} x_{A,n}(\theta_A) dF_{\theta_A} + (1 - \alpha) \int_{\theta_A} x_{A,s}(\theta_A) dF_{\theta_A}$$

C's willingness to pay is analogous to this expression but weights each $\xi \notin \mathfrak{B}_\xi^0$ by θ_A^2, and weights the resulting integral by θ_C^2. That is, we can rewrite C's willingness to pay as

$$WTP_C(\theta_C, \Gamma_A) = \theta_C^2 \Gamma_A$$

where

$$\Gamma_A \equiv \alpha \int_{\theta_A} x_{A,n}(\theta_A)\theta_A^2 dF_{\theta_A} + (1-\alpha) \int_{\theta_A} x_{A,s}(\theta_A)\theta_A^2 dF_{\theta_A}$$

(which is (5)).The function $WTP_C(\theta_C, \Gamma_A)$ is strictly increasing in θ_C, in any equilibrium with $\Gamma_A > 0$. Therefore, if $\Gamma_A > 0$, there exists a unique threshold

$$\theta_C^*(\Gamma_A, p) \equiv \sqrt{\frac{p}{\Gamma_A}} > 0$$

such that C buys m_A if and only if $\theta_C \geq \theta_C^*(\Gamma_A, p)$. If $\Gamma_A = 0$, C's willingness to pay is zero and B would therefore not offer m_A.

We now show that $\theta_C^*(\Gamma_A, p)$ is independent of Γ_A whenever $\Gamma_A > 0$ and p is the monopoly price. This will complete the proof because in equilibrium, p is the monopoly price and $\Gamma_A > 0$ must hold if B is willing to offer m_A (and incur cost c_{B,τ_B}).

For given Γ_A, consider $WTP_C(\theta_C, \Gamma_A)$ as a random variable (driven by randomness in θ_C). Observe that WTP_C has a strictly increasing hazard rate because the hazard rate of θ_C^2 is strictly increasing by assumption.

We now show that θ_C^* is identical for different values of Γ_A that may arise in equilibrium. Towards a contradiction, suppose that there exist two equilibria of the game (equil. (1)) and (equil. (2)) that result in values $(\Gamma_A^{(1)}, \Gamma_A^{(2)})$ with corresponding optimal monopoly prices $(p^{(1)}, p^{(2)})$ and with marginal buyers $(\theta_C^{*(1)} = \theta_C^*(\Gamma_A^{*(1)}, p^{(1)}), \theta_C^{*(2)} = \theta_C^*(\Gamma_A^{(2)}, p^{(2)}))$ such that $\theta_C^{*(1)} > \theta_C^{*(2)}$. That is, we have the following indifference conditions:

$$p^{(1)} = (\theta_C^{*(1)})^2 \Gamma_A^{(1)} \tag{11}$$
$$p^{(2)} = (\theta_C^{*(2)})^2 \Gamma_A^{(2)} \tag{12}$$

Since WTP_C has a strictly increasing hazard rate, we have that $p^{(2)}$ is uniquely optimal for $\Gamma_A^{(2)}$ (see e.g. Börgers, 2015). Thus, in equil. (2) it is strictly better for B to make $\theta_C^{*(2)}$ indifferent than to make $\theta_C^{*(1)}$ indifferent. To achieve the latter, B would have to set a price satisfying

$$\tilde{p} = (\theta_C^{*(1)})^2 \Gamma_A^{(2)} = \frac{\Gamma_A^{(2)}}{\Gamma_A^{(1)}} p^{(1)}$$

where the latter equality follows from plugging in (11). The assumption that $p^{(2)}$ is strictly better than $\tilde{p}$ in equil. (2) applies therefore implies the following revenue inequality:

$$p^{(2)}(1 - F_{WTP_C}^{\Gamma_A^{(2)}}(WTP_C(\theta_C^{*(2)}, \Gamma_A^{(2)}))) > \tilde{p}(1 - F_{WTP_C}^{\Gamma_A^{(2)}}(WTP_C(\theta_C^{*(1)}, \Gamma_A^{(2)}))) \quad (13)$$

$$p^{(2)}(1 - F_{\theta_C}(\theta_C^{*(2)})) > \tilde{p}(1 - F_{\theta_C}(\theta_C^{*(1)})) \quad (14)$$

$$\frac{\Gamma_A^{(1)}}{\Gamma_A^{(2)}} p^{(2)}(1 - F_{\theta_C}(\theta_C^{*(2)})) > p^{(1)}(1 - F_{\theta_C}(\theta_C^{*(1)})) \quad (15)$$

$$\frac{\Gamma_A^{(1)}}{\Gamma_A^{(2)}} p^{(2)}(1 - F_{WTP_C}^{\Gamma_A^{(1)}}(WTP_C(\theta_C^{*(2)}, \Gamma_A^{(1)}))) > p^{(1)}(1 - F_{WTP_C}^{\Gamma_A^{(1)}}(WTP_C(\theta_C^{*(1)}, \Gamma_A^{(1)}))) \quad (16)$$

Now observe that that the LHS of (16) is a feasible revenue in B's maximization problem in equil. (1): multiply both sides of (12) by $\frac{\Gamma_A^{(1)}}{\Gamma_A^{(2)}}$ to see that $\frac{\Gamma_A^{(1)}}{\Gamma_A^{(2)}} p^{(2)}$ is the price that makes $\theta_C^{*(2)}$ indifferent in equil. (1):

$$\frac{\Gamma_A^{(1)}}{\Gamma_A^{(2)}} p^{(2)} = (\theta_C^{*(2)})^2 \Gamma_A^{(1)}$$

All types higher than $\theta_C^{*(2)}$ will therefore also buy and hence the LHS of (16) is the revenue that B gets from setting price $\frac{\Gamma_A^{(1)}}{\Gamma_A^{(2)}} p^{(2)}$. This implies that the RHS of (16) is not the optimal revenue for $\Gamma_A^{(1)}$, contradicting our initial assumption of $p^{(1)}$ being the monopoly price. ∎

Proof of Proposition 2

Like in the main text, we focus on the truncated game played by A and B after having described C's behavior in Proposition 1. The justification for this simplification is recaptured by the observations in this paragraph. In the monopoly game between B and C, they both take A's equilibrium behavior as given, which results in a value of information Γ_A, according to (5). Both types of B optimally set the same price $WTP_C(\theta^*, \Gamma_A)$ if they decide to sell the information, and C follows Proposition 1. As is explained in the main text, the price does not contain information about the message m_A so C's

beliefs, along the equilibrium path, will not be updated upon observing the price. We also have to briefly consider off-equilibrium beliefs. They are unproblematic because the equilibrium must prescribe that if C observes a price that should not arise in equilibrium, he sticks to the strategy to buy if and only if $\theta_C \geq \theta_C^*(\Gamma_A, p)$. This is the uniquely sequentially rational strategy, as otherwise C could be tricked into buying packages that are too expensive for him.

Now fix beliefs of player A about the likelihood of B selling message m_A to C. In equilibrium, the two different cognitive types τ_A have different beliefs, but different θ_A do not. Hence, denote beliefs of type τ_A by $\hat{q}_{B,\tau_A}$ and observe that she maximizes (6),

$$E[\pi_{A,\tau_A}(\theta_A)] = -\theta_A^2(1-\kappa)(1-x_{A,\tau_A}) - \theta_A^2 \kappa \hat{q}_{B,\tau_A} x_{A,\tau_A} \int_{\theta_C^*}^{\bar{\theta}_C} \theta_C^2 dF_C(\theta_C) - c_A(x_{A,\tau_A}),$$

with respect to x_{A,τ_A}. The first two terms of this objective are linear in x_{A,τ_A} and the third term is twice continuously differentiable and concave in x_{A,τ_A}. Hence, $E[\pi_{A,\tau_A}(\theta_A)]$ has a unique maximizer $x_{A,\tau_A}(\theta_A)$. Moreover, at an interior solution the maximizer is continuous and differentiable in $\hat{q}_{B,\tau_A}$, weakly decreases in $\hat{q}_{B,\tau_A}$ and strictly decreases in $\hat{q}_{B,\tau_A}$.

Next consider B's problem. The two types of B simply compare their expected revenue from selling the message, $H(q_{B,n}, q_{B,s})$ to their cost. If

$$H(q_{B,n}, q_{B,s}) \geqslant \hat{c}_{B,\tau_B} \tag{17}$$

holds strictly he chooses $q_{B,\tau_B} = 1$; if it does not hold he chooses $q_{B,\tau_B} = 0$ and if it holds with equality he can mix. Notice that the latter cannot simultaneously be the case for both levels of τ_B because $\tau_B = n$ has a strictly higher cost than $\tau_B = s$. If the low-cost type mixes, the high-cost type does not sell the message. Conversely, if the high-cost type mixes, the low-cost type sells it for sure. It follows that there exist only the five different equilibrium configurations of $(q_{B,n}, q_{B,s})$ that the proposition lists. It remains to show that the equilibrium is unique and that condition (7) determines whether all five, or only four, equilibrium configurations can arise for possible levels of $\hat{c}_{B,\tau_A}$.

To establish uniqueness, we first notice that $H(q_{B,n}, q_{B,s})$ strictly decreases in both arguments if they lie in the open interval $(\underline{q}, \bar{q})$ where $\underline{q}$ is the largest value of $\hat{q}_{B,\tau_A}$ that induces maximal information precision for

all types of A and $\bar{q}$ is the smallest value of $\hat{q}_{B,\tau_A}$ that induces minimal information precision. This strict monotonicity follows from the above observation that the optimal $x_{A,\tau_A}(\theta_A)$ weakly decreases in $\hat{q}_{B,\tau_A}$ and strictly decreases at an interior solution, for all (θ_A, τ_A): an increase in $\hat{q}_{B,\tau_A}$ induces all A types to be more weakly cautious and it induces all A types who provide *some* information to be strictly more cautious. Since the revenue is $H = WTP_C(\theta^*, \Gamma_A)(1 - F_{\theta_C}(\theta^*))$, it follows from (4) and (5) that H decreases strictly in $(q_{B,n}, q_{B,s})$ if all elements of the collection $\{x_{A,\tau_A}(\theta_A)\}_{\theta_A,\tau_A}$ decrease weakly in $(q_{B,n}, q_{B,s})$ and some of them strictly.

We now show that $\underline{q} = 0$. For any possible level of information provision $\tilde{x}_A \geq 0$, the strict concavity of (6) implies that information provision exceeds $\tilde{x}_A$ at belief $\hat{q}_{B,\tau_A}$ iff:

$$\frac{\partial}{\partial x_A}[-\theta_A^2(1-\kappa)(1-x_{A,\tau_A}) - \theta_A^2\kappa\hat{q}_{B,\tau_A}x_{A,\tau_A}\int_{\theta_C^*}^{\bar{\theta}_C}\theta_C^2 dF_C(\theta_C)] \geq c_A'(\tilde{x}_A)$$

$$\theta_A^2(1-\kappa) - \theta_A^2\kappa\hat{q}_{B,\tau_A}\int_{\theta_C^*}^{\bar{\theta}_C}\theta_C^2 dF_C(\theta_C) \geq c_A'(\tilde{x}_A)$$

But for θ_A close to zero, the LHS is close to zero independently of $\hat{q}_{B,\tau_A}$ and of $\tilde{x}_A$, which, together with $c'(\cdot) > 0$ implies that the inequality cannot hold for all θ_A. To induce the largest possible set of θ_A types to provide $\tilde{x}_A$, we need to reduce $\hat{q}_{B,\tau_A}$ and $\tilde{x}_A$ to their minimal levels, which are zero in both cases. That is, $\underline{q} = 0$ and the marginal type θ_A who provides strictly positive information at $\underline{q}$ satisfies $\theta_A^2(1-\kappa) = c_A'(0)$. A type with a large enough θ_A to satisfy this equality exists by assumption.

The upper limit of the named interval, $\bar{q}$, may or may not lie strictly below 1. The case $\bar{q} < 1$ occurs if the most extreme preference type $\bar{\theta}_A$ does not provide any information at belief $\hat{q}_{B,\tau_A} = 1$, i.e. iff

$$\frac{\partial}{\partial x_A}[-\bar{\theta}_A^2(1-\kappa)(1-x_{A,\tau_A}) - \bar{\theta}_A^2\kappa x_{A,\tau_A}\int_{\theta_C^*}^{\bar{\theta}_C}\theta_C^2 dF_C(\theta_C)] < c'(0)$$

$$\bar{\theta}_A^2(1-\kappa) - \bar{\theta}_A^2\kappa\int_{\theta_C^*}^{\bar{\theta}_C}\theta_C^2 dF_C(\theta_C) < c'(0)$$

$$1 - \kappa(1 + \int_{\theta_C^*}^{\bar{\theta}_C}\theta_C^2 dF_C(\theta_C)) < \frac{c'(0)}{\bar{\theta}_A^2},$$

35

which is the negation of condition (7). If $\bar{q} < 1$, then by definition of $\bar{q}$ no type of A reduces x_A any further as $\hat{q}_{B,\tau_A}$ rises above $\bar{q}$. This implies, using (6) and the fact that c_A is twice continuously differentiable, that the optimal $x_{A,\tau_A}(\theta_A)$ cannot be at an interior optimum if $\hat{q}_{B,\tau_A} \geq \bar{q}$ for any θ_A: otherwise, $x_{A,\tau_A}(\theta_A)$ would fall at least a little bit. This implies that $x_{A,\tau_A}(\theta_A) = 0$ holds for all θ_A if $1 \geq \hat{q}_{B,\tau_A} > \bar{q}$. In this case, C's (equilibrium) willingness to pay for messages sent by A players of cognitive type τ_A is zero, too, and thus H is constant in $\hat{q}_{B,\tau_A}$ at $\hat{q}_{B,\tau_A} > \bar{q}$. Moreover, if $1 \geq \hat{q}_{B,\tau_A} > \bar{q}$ holds for both cognitive types τ_A, then $H(q_{B,n}, q_{B,s}) = 0$.

Summing up the previous two paragraphs, we see that $H(q_{B,n}, q_{B,s})$ strictly decreases in both arguments if they lie in $[0, \bar{q})$. $\bar{q}$ may lie strictly below 1, which happens if (7) does not hold. In this case $H(q_{B,n}, q_{B,s})$ is constant in any of it arguments if it lies above $\bar{q}$, and $H(q_{B,n}, q_{B,s}) = 0$ if both arguments lie above $\bar{q}$.

The monotonicity of H implies that the equilibrium is unique. Intuitively, if (17) has slack, then B of type τ_B increases q_{B,τ_B}, and the monotonicity of $H(q_{B,n}, q_{B,s})$ ensures that (17) holds with equality at no more than one combination $(q_{B,n}, q_{B,s})$. Towards a contradiction, suppose that there exist two equilibria $(q_{B,n}^{(1)}, q_{B,s}^{(1)})$, $(q_{B,n}^{(2)}, q_{B,s}^{(2)})$. Different probabilities of τ_B selling can only occur if $H(q_{B,n}, q_{B,s}) = \hat{c}_{B,\tau_B}$ holds in both equilibria. Therefore, both equilibria must yield the same revenue H. With the above monotonicity properties of H, this can arise under exactly two circumstances: either all four probabilities $(q_{B,n}^{(1)}, q_{B,s}^{(1)}, q_{B,n}^{(2)}, q_{B,s}^{(2)})$ lie above $\bar{q}$, or $q_{B,n}^{(i)} < q_{B,n}^{(j)}$ and $q_{B,s}^{(i)} > q_{B,s}^{(j)}$ holds for $i, j \in \{1, 2\}$ and $j \neq i$. In the former case, revenue H is zero and hence all four probabilities must be zero and the two equilibria cannot differ. The latter case implies that both types are indifferent in equilibrium, which is impossible because $\hat{c}_{B,n} > \hat{c}_{B,s}$. This establishes uniqueness.

The borders of the equilibrium regions follow trivially from condition (17). The critical value $H(1, 1)$ is strictly positive iff at least some types of A provide strictly positive information if m_A is sold for sure, i.e. if (7) holds. In this case, it can be that $H(1, 1) \geq \hat{c}_{B,n}$ and the equilibrium is of type I. Regions II, III, IV and V exist for sure (because $H(0, 0)$ and $H(0, 1)$ are strictly positive) and they, respectively, are relevant and correspond to B's behavior as specified in Proposition 2 as follows. II: $H(0, 1) > \hat{c}_{B,n} > H(1, 1)$. III: $\hat{c}_{B,s} \leq H(0, 1)$ and $\hat{c}_{B,n} \geq H(0, 1)$. IV: $H(0, 0) > \hat{c}_{B,s} > H(0, 1)$. V: $H(0, 0) \leq \hat{c}_{B,s}$. ∎

Proposition A.1 In any symmetric equilibrium:

(i) Naive A types provide more: $x_{A,n}(\theta_A) \geq x_{A,s}(\theta_A), \forall \theta_A$

(ii) Interesting A types provide more: $x_{A,\tau_A}(\theta_A)$ increases weakly in $|\theta_A|, \forall \tau_A$

Proof Both results are almost immediate from A's objective function in (6). Clause (i) uses the observation that $\hat{q}_{B,n} \leq \hat{q}_{B,s}$ (see Proof of Proposition 2), which implies the result together with (6). Clause (ii) follows from (6) by observing that A's benefit from transmitting information is linear in x_A and quadratic in θ_A: the optimal information provision is 0 if

$$1 - \kappa(1 + \hat{q}_{B,\tau_A} \int_{\theta_C^*}^{\bar{\theta}_C} \theta_C^2 dF_C(\theta_C)) < \frac{c'(0)}{\theta_A^2},$$

and otherwise it increases strictly in θ_A^2. ∎

We now add a proposition about the players' (expected) equilibrium payoffs as we vary $(\hat{c}_{B,n}, \hat{c}_{B,s})$. (We drop arguments and expectation operators for brevity.) In order to accommodate the two interpretations for player B (see the introduction), we split his payoff into two parts, π_{B_1} resulting from the direct interaction with player A and π_{B_2} resulting from the sale of the message to player C.

Proposition A.2

(i) $\frac{d\pi_{A,s}}{d\hat{c}_{B,s}} \geqslant 0$ and $\frac{d\pi_{A,s}}{d\hat{c}_{B,n}} = 0$

(ii) $\frac{d\pi_{A,n}}{d\hat{c}_{B,s}} \geqslant 0$ and $\frac{d\pi_{A,n}}{d\hat{c}_{B,n}} \leqslant 0$

(iii) $\frac{d\pi_{B_1}}{d\hat{c}_{B,s}} \geqslant 0$ and $\frac{d\pi_{B_1}}{d\hat{c}_{B,n}} \geqslant 0$

(iv) $\frac{d\pi_{B_2}}{d\hat{c}_{B,s}} \leqslant 0$ and $\frac{d\pi_{B_2}}{d\hat{c}_{B,n}} \geqslant 0$

(v) $\frac{d\pi_C}{d\hat{c}_{B,s}} = 0$ for $\hat{c}_{B,s} \leqslant H(0,1)$; $\pi_C = 0$ for $\hat{c}_{B,s} \geqslant H(0,0)$; $\frac{d\pi_C}{d\hat{c}_{B,s}}$ ambiguous otherwise and $\frac{d\pi_C}{d\hat{c}_{B,n}} \geqslant 0$

Proof For the proof of the proposition, we first introduce two Lemmas.

Lemma 1 $\frac{dq_{B,\tau_B}}{d\hat{c}_{B,\tau_B}} \leqslant 0$ and $\frac{dq_{B,\tau_B}}{d\hat{c}_{B,\tau \neq \tau_B}} = 0$, for all τ_B.

Proof of Lemma 1 Let us first examine the first part of the lemma. Take $\tau_B = n$. For $\hat{c}_{B,s} \geqslant H(0,1)$ the high-cost B never sells the message. For $\hat{c}_{B,s} < H(0,1)$ we know that as we increase $\hat{c}_{B,n}$ from zero upwards we are first in equilibrium region I where the high-cost B always sells, then in region II where he mixes, and finally in region III where he never sells. Hence, the proof will be complete if we can show that the claim holds *within* region II. In this region, the high-cost B is indifferent, i.e. $H(q_{B,n}, q_{B,s}) = \hat{c}_{B,n}$. Recall from the Proof of Proposition 2 that H is monotonically decreasing in both arguments, which yields the result for $\tau_B = n$. A similar logic applies to $\tau_B = s$: $q_{B,s}$ varies with $\hat{c}_{B,s}$ only in region IV, where the monotonicity of H and the low-cost B's indifference yield the result. For the second part of the lemma, notice that the cost of the other type matter only insofar as they change the information provision by the type of player A who believes in the other cost level. Let us examine $\tau_B = s$. For $\hat{c}_{B,s} < H(0,1)$ the low-cost B always sells the message: variations in information provision by the naive A are irrelevant. For larger $\hat{c}_{B,s}$, notice that due to $q_{B,n}$ being constant there is no change in information provision by the naive A who believes in high costs. The same logic applies to $\tau_B = n$. ∎

The next lemma shows that for an arbitrary change in parameters, the payoff of sophisticated A and the profits that B obtains from his direct interaction with A move in parallel with (i.e. their difference has the same sign as the difference in) sophisticated A's information precision

Lemma 2 For all θ_A and for any parameter γ that is either in $\{\kappa, \alpha, \hat{c}_{B,n}, \hat{c}_{B,s}\}$ or is a distribution parameter of F_A or F_C (but not a parameter of A's cost function): $sign(d\pi_{A,s}(\theta_A)/d\gamma) = sign(d\pi_{B_1}/d\gamma) = sign(dx_{A,s}(\theta_A)/d\gamma)$.

Proof of Lemma 2 To see that $sign(d\pi_{B_1}/d\gamma) = sign(dx_{A,s}(\theta_A)/d\gamma)$, observe from (6) that the information provision of sophisticated A moves in the same direction for all θ_A, i.e. $sign(dx_{A,s}(\theta_A)/d\gamma) = sign(dx_{A,s}(\theta'_A)/d\gamma)$ for all (θ_A, θ'_A). Therefore, B benefits more from his interaction with *sophisticated* A iff $dx_{A,s}(\theta_A)/d\gamma \geq 0$. For the result to be wrong, we would need that *naive* A strictly adjusts her information provision in the opposite direction. But in equilibrium, the beliefs about B need to describe B's selling behavior as optimal according to (17). There cannot be a change in H or in one of the cost parameters $(\hat{c}_{B,n}, \hat{c}_{B,s})$ that would strictly add slack to (17) for one type of B and strictly reduce slack for the other type. Hence (using (6) again)

there cannot exist a change in γ that induces a strict increase in $x_{A,s}$ but a strict reduction in $x_{A,n}$, or vice versa.

To see that $sign(d\pi_{A,s}(\theta_A)/d\gamma) = sign(dx_{A,s}(\theta_A)/d\gamma)$, consider sophisticated A's expected payoff in (6). Its first two parts are linear in $x_{A,s}$ and the convex cost of $x_{A,s}$ is subtracted. Hence, A's payoff equals the area between a constant line (reflecting the marginal benefits of transmitting information) and some increasing function capturing A's marginal costs. The latter remains unaffected by γ. So any change in γ will only move the constant marginal benefit. (6) implies that if γ moves that marginal benefit upward (downward) both the optimal $x_{A,s}$ and A's profits go up (down). ∎

We are still in the proof of Proposition A.2. Part (i) and part (iii) follow immediately from combining the Lemma 1 and Lemma 2. For the first half of part (ii) notice that whatever the naive A actually does, increases in the actual costs make it according to Lemma 1 weakly less likely that his information will actually be sold to C. For the second half part (ii) notice that for $\hat{c}_{B,s} \leqslant H(0,1)$ the information is always sold and that according to Lemma 1 higher imagined costs make the naive A less cautious such that his true profits fall. For $\hat{c}_{B,s} > H(0,1)$ naive A's behavior is constant in $\hat{c}_{B,n}$ and changes in imagined costs have no further effect on his actual payoff. For the first half of part (iv) notice first that B earns zero from the information sale in equilibrium regions IV and V, that is, when $\hat{c}_{B,s} \geqslant H(0,1)$. For lower $\hat{c}_{B,s}$ notice that his actual costs do not affect the amount of information that is provided. However, he obviously has to pay $c_{B,s}$. For the second half of part (iv) use Lemma 1 and (6) to see that A provides more information in response to an increase in $\hat{c}_{B,n}$, and notice that B does not actually pay the imagined cost $\hat{c}_{B,n}$, while his revenue increases. Finally, for part (v) observe that in equilibrium regions I, II, and III changes in $\hat{c}_{B,s}$ neither affect the amount of information that players A will provide nor the probability that the true (low-cost) B offers m_A for sale. In region V, B never sells, hence $\pi_C = 0$. In region IV, increasing costs imply that B sells with a decreasing probability. This increases the amount of information that the sophisticated A provides and hence C's expected surplus conditional on buying. However, at the same time it becomes less likely that that surplus will materialize. Finally, for the last statement, $\frac{d\pi_C}{d\hat{c}_{B,n}} \geqslant 0$, observe that for $\hat{c}_{B,s} < H(0,1)$ higher imagined costs provide C with more information from naive A while the information provision from sophisticated A as well as B's probability of selling remain constant, and that if $\hat{c}_{B,s} \geq H(0,1)$, both are constant in $\hat{c}_{B,n}$. ∎