Correlation-Savvy Sellers

Roland Strausz (HU Berlin)

Discussion Paper No. 347

November 10, 2022

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Georg Weizsäcker, Ph.D., Humboldt University Berlin, 10117 Berlin, Germany
info@rationality-and-competition.de

Correlation-Savvy Sellers

Roland Strausz*

November 5, 2022

Abstract

A multi-product monopolist sells sequentially to a buyer who privately learns his valuations. Using big data, the monopolist learns the intertemporal correlation of the buyer's valuations. Perfect price discrimination is generally unattainable—even when the seller learns the correlation perfectly, has full commitment, and in the limit where the consumption good about which the buyer has ex ante private information becomes insignificant. This impossibility is due to informational externalities which requires information rents for the buyer's later consumption. These rents induce upward and downward distortions, violating the generalized no distortion at the top principle of dynamic mechanism design.

*Humboldt-Universität zu Berlin, Institute for Economic Theory 1, Spandauer Str. 1, D-10178 Berlin (Germany), Email: strauszr@wiwi.hu-berlin.de. I thank Helmut Bester, Alex Gershkov, Amir Habibi, Maren Hahnen, Daniel Krähmer, and seminar participants at the MACCI conference in Annweiler, Toulouse School of Economics, and Peking University for very helpful comments and discussions. Financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190 (project number 280092119) is gratefully acknowledged.

1 Introduction

In recent decades, the tracking of consumers has become standard practise in both online and offline consumer markets. In online markets, retailers not only track their customers as soon as they log into their personal accounts, they also track them through the use of cookies when they are not logged in. Brick and mortar stores track their consumers by linking scanner data to credit card payments and by setting up loyalty or membership programs, which, in terms of tracking, play a similar role to the consumer's online account at an online shop. All this tracking yields massive amounts of data—big data, through which retailers sift continuously with the help of data scientists.

While data science is hailed for revolutionizing consumer markets by the use of highly sophisticated computing techniques, boiled down to its essence, it tries to identify robust correlations that are potentially valuable to retailers. One of the first and by now classical example is the detection by a data scientist in 1992 of an unexpected positive correlation between the evening sales of beer and diapers, prompting the retailer to group together the two products, and thereby raising sales.[1] While novel at the time, it is now standard that based on big data, online platforms such as Amazon and Netflix make personalized suggestions for buying products and viewing movies.[2] It has also become more and more prevalent to use this information for sending out personalized vouchers, thereby allowing retailers to price discriminate.

Yet, even though the high investments of retailers into data science techniques signify their practical importance, our basic economic understanding of sellers learning such correlations is still limited. Indeed, most of our insights about the impact of big data are based on models that consider sellers who learn directly the private information of a particular consumer. However, it is unclear whether such models effectively capture the idea that tracking and big data allow firms to learn about the correlations of a buyer's preferences between different goods rather than his preferences directly.

To identify such possible discrepancies, I consider a setup of a "correlation-savvy" seller, who uses big data to learn only about correlations. Its analysis confirms that learning about correlation differs from learning about preferences directly. When big data allows the seller to learn directly and perfectly the buyer's preferences, it is immediate that she can extract all consumption rents and has no incentive to distort future allocations. By contrast, when big data allows the monopolist to learn only about the correlation, then, in general, this does not enable her to extract fully the buyer's consumption rents; even in the best possible case, where the monopolist learns the correlation perfectly and has full commitment ex ante. Consequently, the seller has an incentive to distort also the future allocations.

I illustrate this result in a two period model with binary valuations and a correlation structure where types are either perfectly positively correlated—they are persistent—or perfectly negatively correlated—they switch. I show that, in general, a correlation-savvy seller who has full commitment and perfectly learns the correlation structure cannot extract all the buyer's consumption rents in the second period. The analysis reveals that this inability is due to the informational externality that information

[1] Source: Mark Madsen (2017), "Beer, diapers, and correlation: A tale of ambiguity", keynote address ITWeb Business Intelligence Summit 2017, `http://download.1105media.com/tdwi/Remote-assets/Events/2017/Boston/MarkMadsen_Beer-and-Diapers.pdf` (last retrieved 16.11.2021).

[2] This indicates that big data enable sellers to get to know the tastes of their customers better than the consumers themselves. See Hariri (2018) for an elaborate discussion of the fact that data science allows to obtain better knowledge about individuals than the individuals have themselves.

about the buyer's ex ante valuation and the correlation structure are necessarily complementary. In particular, the seller can learn the buyer's future valuation only if she knows both the correlation structure and the buyer's initial type. This reveals a fundamental difference between learning about correlations and learning about types directly.

The two-period, binary valuation model is tractable enough to completely characterize the optimal full commitment contract for all possible parameter constellations. It therefore also allows a full characterization of its comparative statics. Moreover, the modelling setup allows to argue that its results are robust to the exact timing when the seller learns about the correlation structure, to whether she observes it privately or publicly, and to whether she learns the correlation only imperfectly.

2 Related Literature

The current paper belongs to the literature of behavioral based price discrimination which considers optimal pricing policies of monopolists, who are able to learn about the consumers' purchase history and/or their personal tastes (e.g., Fudenberg and Villas-Boas, 2006). In particular, the setup coincides with a two period-version of the sequential selling problem of Battaglini (2005) but with the extension that the seller obtains additional information about the transition matrix of the buyer's valuation. One of the main insights of Battaglini (2005) is the *generalized no distortions at the top principle*: Allocations are distorted only for histories where the buyer never had a high valuation. I show that when the seller learns about correlations, this principle no longer holds. The reason for this failure is that different dynamic incentive constraints are binding at optimum, leading to different economic distortions.

As emphasized in the introduction, learning in my setup differs from studies that consider monopolistic sellers that learn about a consumer's private information directly and use this information to price discriminate (e.g., Conitzer et al. 2012, Bergemann et al. 2015, De Cornière and De Nijs 2016, Ali et al. 2019, Bonatti and Cisternas 2020, De Cornière and Taylor 2020). In contrast to the current paper, perfect learning in this literature allows a monopolist to extract all information rents. By studying a different type of learning, the analysis complements this literature.

Contrasting complementarities between the seller's and the buyer's private information as studied here, a recent literature studies data externalities (e.g., Choi et al. 2019, Acemoglu et al. 2019. Bergemann et al. 2019, Ichihashi 2021). Because they reflect informational interdependencies between the consumers themselves, these externalities are orthogonal to the externalities that I study. Börgers et al. (2013) provide a more fundamental analysis of informational externalities.

Finally, this paper belongs to a nascent literature which emphasizes that big data analysis allows firms to gain an informational advantage over consumers, leading to an "inverse" screening problem in which firms have private information about their customer's preferences. In an insurance setup with two dimensional types, Brunnermeier et al. (2021) consider an insurer who obtains private information about the correlation of a consumer's two dimensional risk type. In contrast to the current paper, the authors study the effects of these correlation savvy insurers on boundedly rational consumers and competition. Ichihashi and Smolin (2022) study a seller, who receives a private signal that is informative about the buyer's valuation. In contrast to the current paper, the seller learns about individual types directly rather than about correlation structures. In a spatial model of bank competition, Vives and Ye

(2021) study how advances in informational technology enable banks to know more about the success probability of projects than the entrepreneurs to whom they provide finance. In a similar vein but abstracting from competition, Strausz (2009) considers investors who are better informed about the success probability of projects than entrepreneurs, focusing on the extent to which optimal contracts induce the investor to reveal this private information to the entrepreneur.

3 The Setup

Consider a seller (she), who first provides a quantity of some good q to a buyer (he), and subsequently some quantity Q of some other good. The transaction involves an overall transfer $T \in \mathbb{R}$ from the buyer to the seller. Hence, the economic allocation is a triple $(T, q, Q) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+$. I am agnostic about the physical relationship between the goods q and Q. That is, they can represent the same good so that the model is one of repeated purchases, or Q may pertain to goods that are physically unrelated to good q, such as beer and diapers.

Payoffs. Measuring the importance of good Q relative to good q by a parameter $\delta \geq 0$, the seller's profit and the buyer's utility associated with an allocation (T, q, Q) are, respectively,

$$\Pi(T, q, Q) = T - c(q) - \delta C(Q) \quad \text{and} \quad U(T, q, Q|\theta, \Theta) = -T + \theta q + \delta \Theta Q,$$

where $c(.)$ and $C(.)$ represent the seller's cost functions, and (θ, Θ) the agent's marginal valuation for quantities q and Q, respectively. I assume that the cost functions are twice differentiable, increasing, convex, and exhibit $c(0) = c'(0) = C(0) = C'(0) = 0$ and $\lim_{q\to\infty} c(q) = \lim_{q\to\infty} c'(q) = \lim_{Q\to\infty} C(Q) = \lim_{Q\to\infty} C'(Q) = \infty$. Concerning the valuation (θ, Θ), I assume that they are binary; they can either be high, h, or low, l, with $\Delta \equiv h - l > 0$.

Efficient quantities equalize marginal costs to marginal utility, and, hence, the efficient quantities $(q_l^*, q_h^*, Q_l^*, Q_h^*)$ satisfy

$$c'(q_h^*) = h; \; c'(q_l^*) = l; \; C'(Q_h^*) = h; \; C'(Q_l^*) = l.$$

The assumptions on the cost functions together with $\Delta > 0$ imply $q_h^* > q_l^* > 0$ and $Q_h^* > Q_l^* > 0$.

Information structure. Initially, the buyer privately knows his valuation of the initial good q, whereas the seller only knows that this valuation is high with probability $\mathbb{P}\{\theta = h\} = \nu$ and low with probability $\mathbb{P}\{\theta = l\} = 1 - \nu$. I focus on the case that the correlation between the valuations is either perfectly positive or perfectly negative. Hence, the correlation structure $\gamma \in \{p, s\}$ is such that the buyer's valuation is either persistent, $\gamma = p$, or switches, $\gamma = s$. Ex ante, there is no private information about the correlation; the buyer and seller commonly know that the buyer's value type is persistent with probability $\mathbb{P}\{\Theta = \theta\} = \pi$, and switches with probability $\mathbb{P}\{\Theta \neq \theta\} = 1 - \pi$. The draw of the initial type and its persistence are stochastically independent so that the joint probability distribution of (θ, Θ) is

$$\mathbb{P}\{(\theta, \Theta) = (h, h)\} = \nu\pi; \qquad \mathbb{P}\{(\theta, \Theta) = (h, l)\} = \nu(1 - \pi);$$
$$\mathbb{P}\{(\theta, \Theta) = (l, l)\} = (1 - \nu)\pi; \qquad \mathbb{P}\{(\theta, \Theta) = (l, h)\} = (1 - \nu)(1 - \pi).$$

Timing. The buyer privately learns his valuation θ before the contracting stage, and privately learns valuation Θ after consuming q. The seller privately learns the correlation structure γ after the contracting stage but before the buyer consumes the good q. Hence, the timing of events is as follows:

0. The buyer privately learns $\theta \in \{l, h\}$;
1. Seller offers a contract determining the terms of trade (T, q, Q);
2. Buyer decides whether to reject or accept the contract;
3. Seller privately learns the correlation structure $\gamma \in \{p, s\}$;
4. Buyer consumes q;
5. Buyer learns $\Theta \in \{l, h\}$;
6. Payoffs are realized from consuming quantities (q, Q) and transfer T.

Remarks. Before analyzing the seller's profit-maximizing outcome in this setup, it is helpful to discuss its relation to the paper's motivating example in the introduction and the structure of the contracting terms (T, q, Q).

First, consider the motivating diaper and beer example as discussed in the introduction. One ex post rationalization for the positive correlation between the evening sales of these products is the following.[3] At the end of a working day, a young father receives a telephone call from his spouse that they have run out of diapers, thereby learning that $\theta = h$. When picking up the diapers in the shop, the father runs into the beer stand and discovers his high willingness to pay for beer, learning $\Theta = h$. The model's sequential timing of the buyer learning first θ and subsequently Θ reflects this ex post rationalization.[4]

Moreover, the assumption that the seller learns about the buyer's persistence after the contracting stage reflects a setting in which the seller learns about the buyer's correlation from combining her big data with the buyer's specific characteristics (which she receives after the contracting stage). While this is one possible view of big data analysis, an alternative view is that big data allows the seller to learn about the consumer's persistence without any personal data of the consumer. In this case, the seller would learn about the persistence of the consumer's valuation before the contracting stage and this seems more in line with the motivating beer and diaper example. Usually the exact timing of receiving private information crucially affects equilibrium outcomes. This is however not the case here; the seller's profit-maximizing equilibrium outcome that obtains with the specific timing above remains profit-maximizing when the seller learns the correlation before offering the contract. In fact, the outcome also obtains when the seller learns it only after the consumption of quantity q. Hence, the profit-maximizing equilibrium outcome is robust with respect to the point in time at which the seller learns the correlation, and the specific timing above allows me to demonstrate this robustness.

Second, the description above leaves open what kind of contracts the seller can offer to the buyer. For multiple reasons, it is instructive to allow the seller to offer any possible contract rather than restrict her contracting space artificially. Firstly, this complete contracting approach represents the best case-scenario for the seller. Hence, if the seller cannot attain the perfect price discrimination outcome under this scenario, then she can also not attain it when her contracting possibilities are more limited. Complete contracting therefore ensures that any inability of the seller to attain perfect price discrimination

[3] See for instance `https://canworksmart.com/diapers-beer-retail-predictive-analytics/`

[4] Note that if the father had already known his valuation for beer before entering the shop, the shop could not have raised its revenue by positioning the beer stand close to the diapers.

is due to her learning about correlations rather than some implicit incomplete contracting assumption. Secondly, studying complete contracts allows a direct comparison to related classical models in the literature and, thereby, allows to pin down the economic effects of sellers who learn about correlation structures. Indeed, without the seller learning the correlation in stage 3, the model reduces to a two-period version of the dynamic mechanism design problem of Battaglini (2005). Hence, this setup is a natural benchmark for identifying the economics effects of correlation-savvy sellers and I explicitly introduce it in Section 4. Thirdly, taking a complete contracting approach yields a tractable model, because, as argued in Section 5, it is amendable to a dynamic revelation principle. Finally, the analysis provides a normative benchmark of what a correlation-savvy seller should do if she wants to take full advantage of learning about correlations.

These reasons motivate my study of correlation-savvy sellers by taking a complete contracting approach. The approach presumes that the seller can fully commit to any long term contract that determines the terms trade (T, q, Q).

4 Benchmarks

Before deriving the profit-maximizing long term contract explicitly, it is helpful to introduce first three relevant benchmarks to which to compare my results and thereby identify the economic effects of a correlation-savvy seller: the first best, a static framework with private information, and a dynamic setup but without the learning stage 3.

4.1 Perfect Price Discrimination Benchmark

First consider the outcome under perfect information, where the seller directly observes the agent's overall type (θ, Θ). In this case, the seller can price discriminate perfectly and extract the whole surplus by offering type (θ, Θ) the type-specific efficient quantities q^*_θ and Q^*_Θ for an overall transfer of $T_{\theta\Theta} = \theta q^*_\theta + \delta \Theta Q^*_\Theta$. I refer to this outcome of efficient quantities and full surplus extraction by the seller as the perfect price discrimination outcome.

4.2 Static Benchmark

As a second benchmark, it is useful to consider a pure static version of the model in which learning about the correlation structure does not matter. Such a model obtains for $\delta = 0$, where the model boils down to a static, single-good monopolistic screening problem in the tradition of Mussa and Rosen (1978) with binary types. Recall that in this classic setup, it is optimal for the seller to offer a direct mechanism $\{(q_l, p_l), (q_h, p_h)\}$ that maximizes her expected profits

$$\Pi = \nu(p_h - c(q_h)) + (1 - \nu)(p_l - c(q_l))$$

subject to the incentive compatibility constraints

$$hq_h - p_h \geq hq_l - p_l \text{ and } lq_l - p_l \geq lq_h - p_h; \tag{1}$$

and the individual rationality conditions

$$hq_h - p_h \geq 0 \text{ and } lq_l - p_l \geq 0. \tag{2}$$

As is well known, the seller's first best outcome of selling the efficient quantity q^*_θ to each type θ at a respective price θq^*_θ, is not attainable as it violates the incentive constraint of the efficient type h. Indeed, at the optimum, the incentive constraint of the efficient and the individual rationality of the inefficient type constrain the seller and induces her to distort the quantity q_l downwards. In particular, it leads the seller to offer type l a quantity $q^h_l < q^*_l$ such that[5]

$$c'(q^h_l) = (l - \varphi\Delta)^+,$$

where $\varphi \equiv \nu/(1-\nu)$ is the relative likelihood of an efficient type. As a future point of reference, the following proposition summarizes the outcome in this benchmark.

Proposition 0 *Suppose $\delta = 0$. Then it is optimal for the seller to offer the incentive compatible menu $\{(q_l, p_l), (q_h, p_h)\}$ with the efficient quantity $q_h = q^*_h$ for type h and a downward distorted quantity $q_l = q^h_l < q^*_l$ for type l, leaving a positive rent to type h and no rents to type l.*

4.3 Dynamic Mechanism Design Benchmark

Without the seller learning the correlation in stage 3, the model boils down to a two-period version of Battaglini (2005). For this benchmark, the dynamic revelation principle in Myerson (1986) implies that one can express the profit-maximizing contract as an incentive compatible direct mechanism that induces the buyer to first report truthfully his valuation θ, and, after learning Θ, to report truthfully his subsequent valuation Θ. Consequently, it is without loss to focus on menus $(T_{\tilde{\theta}\tilde{\Theta}}, q_{\tilde{\theta}}, Q_{\tilde{\theta}\tilde{\Theta}})$ that condition the terms of trade (T, q, Q) on an initial report $\tilde{\theta}$ about θ and a subsequent report $\tilde{\Theta}$ about Θ, and that satisfy individual rationality constraints and dynamic incentive constraints which induce honest reporting on the equilibrium path.

For type $\theta = h$, such honest reporting requires that for all $i, j, k \in \{h, l\}$, it holds

$$hq_h + \pi(\delta h Q_{hh} - T_{hh}) + (1-\pi)(\delta l Q_{hl} - T_{hl}) \geq hq_i + \pi(\delta h Q_{ij} - T_{ij}) + (1-\pi)(\delta l Q_{ik} - T_{ik}). \tag{3}$$

Hence, for type $\theta = h$ there are effectively $2^3 - 1 = 7$ dynamic incentive constraints. They reflect the different ways in which the type can combine misreports about θ and Θ.

For type $\theta = l$, there are also 7 dynamic incentive constraints: for all $i, j, k \in \{h, l\}$, it must hold

$$lq_l + \pi(\delta l Q_{ll} - T_{ll}) + (1-\pi)(\delta h Q_{lh} - T_{lh}) \geq lq_i + \pi(\delta l Q_{ij} - T_{ij}) + (1-\pi)(\delta h Q_{ik} - T_{ik}). \tag{4}$$

Battaglini (2005) focuses on cases in which the distribution of the valuation Θ for type $\theta = h$ stochastically dominates this distribution for type $\theta = l$. In my setup, such stochastic dominance is equivalent to assuming $\pi > 1/2$. For this parameter constellation, it follows from Battaglini (2005) that of the 14

[5] To deal with corner solutions, let $(x)^+$ denote the positive part of a number x, i.e. $(x)^+ \equiv \max\{0, x\}$.

dynamic incentive constraints only two incentive constraints constrain the optimum. First, the profit-maximizing contract is restricted by the need to dissuade type $\theta = h$ to claim to be type $\theta = l$ together with claiming that his type Θ remains l. Second, it is restricted by the need to dissuade type $\theta = l$ to claim type $\Theta = l$ in the case that type $\theta = l$ switches into type $\Theta = h$.

Maximizing the seller's profits under these two constraints then yields that, except for the allocations q_l and Q_{ll}, all allocations are efficient. Battaglini (2005) calls this the *generalized no distortion at the top (GNDT) principle*.

5 The Seller's Optimization Problem

In order to state the seller's problem of finding the profit maximizing contract as a tractable optimization problem, I first derive in this section a specific class of mechanisms to which the seller's profit-maximizing contract belongs.

Direct Mechanisms. The starting point is Myerson (1986), whose results imply that in the framework of this paper, the seller's profit-maximizing contract cannot outperform a direct mechanism with the following properties: 1) it requests confidential reports from the buyer and seller about their private information as soon as they receive it, and ii) it ensures that each player has an incentive to report truthfully given that a player expects the other to report also truthfully.

This suggests to focus on direct mechanisms of the form

$$\phi = (T_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}, q_{\tilde{\theta}\tilde{\gamma}}, Q_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}),$$

which condition the economic allocation (T, q, Q) on the buyer's reports $\tilde{\theta} \in \{l, h\}$ about θ, the seller's report $\tilde{\gamma} \in \{p, s\}$ about the correlation structure (i.e., whether the buyer's valuation is persistent, $\gamma = p$, or switches, $\gamma = s$), and the buyer's report $\tilde{\Theta} \in \{l, h\}$ about Θ. Because the buyer learns Θ only after consuming the quantity q, the quantity q cannot be conditioned on the report $\tilde{\Theta}$.

A potential problem is however that the required confidentiality of the seller's report $\tilde{\gamma}$ may stand in conflict with the execution of the direct mechanism. In particular, if the direct mechanism ϕ conditions the quantity q non-trivially on $\tilde{\gamma}$, then the buyer can deduce the seller's report $\tilde{\gamma}$ from her consumption of q before reporting Θ. This would then undermine the presumed confidentiality of $\tilde{\gamma}$. My solution to this problem is to side-step this issue completely (i.e., implicitly assuming that such deductions do not take place), and show that, in fact, the seller's profit-maximizing direct mechanism does not condition the quantity q on the report $\tilde{\gamma}$, implying that deducing $\tilde{\gamma}$ from q is indeed not an issue.

Summarizing, a direct mechanism ϕ induces the following game. First, the buyer learns his ex ante type θ about which he sends a report $\tilde{\theta} \in \{l, h\}$. Subsequently the seller learns the correlation γ about which she sends a report $\tilde{\gamma} \in \{p, s\}$. Finally, the buyer learns his ex post type Θ about which he sends a report $\tilde{\Theta} \in \{l, h\}$, while being ignorant of the seller's report $\tilde{\gamma}$. The outcome of the game is an allocation $(T_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}, q_{\tilde{\theta}\tilde{\gamma}}, Q_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}})$, inducing the respective terms of trade that determine the payoffs of the seller and buyer.

Incentive Compatibility. I next address the issue of incentive compatibility, raising the question which direct mechanisms ϕ induce a game as described above in which truthful reporting is an equilibrium.

A first observation is that, by exploiting the correlation between γ and Θ, the seller can costlessy ensure that she reports honestly. To see this, note that if both the seller and the buyer report truthfully, the following sequences of reports $(\tilde{\theta}, \tilde{\gamma}, \tilde{\Theta})$ never occur on the equilibrium path:

$$\mathcal{O} \equiv \{(l,s,l),(l,p,h),(h,s,h),(h,p,l)\}.$$

These out-of-equilibrium events imply direct mechanisms which induce the seller to report γ honestly without the need for any information rents. To see this, note that if the seller expects the buyer to report truthfully, then she has an incentive to report the persistent correlation structure, $\gamma = p$, truthfully if

$$\nu\Pi(T_{hph}, q_{hp}, Q_{hph}) + (1-\nu)\Pi(T_{lpl}, q_{lp}, Q_{lpl}) \geq \nu\Pi(T_{hsh}, q_{hs}, Q_{hsh}) + (1-\nu)\Pi(T_{lsl}, q_{ls}, Q_{lsl}).$$

Hence, by setting the out-of-equilibrium quantities Q_{hsh} and Q_{lsl} large enough, the seller reports truthfully. Likewise, the seller has an incentive to report the switching correlation structure, $\gamma = s$, truthfully if the out-of-equilibrium quantities Q_{hpl} and Q_{lph} are large enough.

Similarly, setting the out-of-equilibrium transfers T_{lsl}, T_{lph}, T_{hsh}, T_{hpl} large enough ensures that, given a truthful report θ, the buyer also reports his private information Θ truthfully. Hence, by punishing both the buyer and the seller for an out-of-equilibrium reporting triple in $\mathcal{O}$, an honest reporting of the correlation structure γ and the valuation Θ does not require any information rents..[6,7]

It follows that the only relevant incentive constraints are the ones that prevent the buyer from misreporting outcomes that are not in $\mathcal{O}$. To formalize these incentive constraints, note that the buyer's action is a report $\tilde{\theta} \in \{h, l\}$ about θ and a report $\tilde{\Theta} \in \{h, l\}$ about Θ. Because the buyer can perfectly deduce the seller's observation γ from (θ, Θ), the buyer can effectively condition his report $\tilde{\Theta}$ on the seller's observation of γ. This means that the buyer's reporting strategy is a triple $r = (r_1, r_2, r_3) \in S \equiv \{l,h\} \times \{l,h\} \times \{l,h\}$, where r_1 represents the report $\tilde{\theta}$, r_2 represents the report $\tilde{\Theta}$ given the switching correlation structure $\gamma = s$, and r_3 represents the report $\tilde{\Theta}$ given the persistent correlation structure $\gamma = p$. As the buyer expects the seller to report her observed correlation truthfully, the expected payoff of an ex ante buyer type θ who uses a reporting strategy $r = (r_1, r_2, r_3)$ is

$$U(r_1, r_2, r_3|\theta) = \pi[q_{r_1p}\theta + \delta Q_{r_1pr_2}\theta - T_{r_1pr_2}] + (1-\pi)[q_{r_1s}\theta + \delta Q_{r_1sr_3}\theta' - T_{r_1sr_3}],$$

where θ' is the singleton of the set $\{h,l\}\backslash\{\theta\}$.

For an ex ante type $\theta = h$, honest reporting means to pick the triple $(r_1, r_2, r_3) = (h, l, h)$ rather than $(l, h, l) \notin \mathcal{O}$. Hence, truthtelling requires $U(h,l,h|h) \geq U(l,h,l|h)$, yielding the incentive constraint:

$$\begin{aligned} &\pi[q_{hp}h + \delta Q_{hph}h - T_{hph}] + (1-\pi)[q_{hs}h + \delta Q_{hsl}l - T_{hsl}] \geq \\ &\quad \pi[q_{lp}h + \delta Q_{lpl}h - T_{lpl}] + (1-\pi)[q_{ls}h + \delta Q_{lsh}l - T_{lsh}]. \end{aligned} \tag{5}$$

[6]As formally shown in the proof of Lemma 1, quantities and transfers exist which, conditional on an honest report $\tilde{\theta}$, induces both the seller and the buyer to report their private information honestly without the need of additional information rents.

[7]As shown explicitly in Section 8, the presence of out-of-equilibrium events is not crucial for the result that an elicitation of the correlation structure γ and the valuation Θ requires no information rents. The reason is that types γ and Θ are inherently correlated so that, in the spirit of Crémer and McLean (1985), the use of transfers and lotteries enables a costless extraction of private information.

For an ex ante type $\theta = l$, honest reporting means to pick the triple $(r_1, r_2, r_3) = (l, h, l)$ rather than $(h, l, h) \notin \mathscr{O}$. Hence, truthtelling requires $U(l, h, l|l) \geq U(h, l, h|l)$, yielding the incentive constraint:

$$\begin{aligned}
&\pi[q_{lp}l + \delta Q_{lpl}l - T_{lpl}] + (1-\pi)[q_{ls}l + \delta Q_{lsh}h - T_{lsh}] \geq \\
&\pi[q_{hp}l + \delta Q_{hph}l - T_{hph}] + (1-\pi)[q_{hs}l + \delta Q_{hsl}h - T_{hsl}].
\end{aligned} \tag{6}$$

It is instructive to compare these incentive constraints to the benchmark of Section 4.3, where the seller does not learn the correlation structure. Recall that constraints (5) and (6) dissuade type θ to misreport both θ and Θ. These two particular constraints are also part of the 14 dynamic incentive constraints that obtain in the dynamic mechanism design benchmark of Section 4.3. However, as argued, the results of Battaglini (2005) imply that, for $\pi > 1/2$, these two particular constraints do not restrict the seller's optimum. In fact, the relevant constraints in Section 4.3 are the one that dissuades type $\theta = h$ to claim to be type $\theta = l$ together with claiming his type remains l, and the one that dissuades type $\theta = l$ to claim type $\Theta = l$ in the case that type $\theta = l$ switches into type $\Theta = h$. The key observation is that when the seller learns about the correlation structure, these two ways of misreporting lead to some out-of-equilibrium report in $\mathscr{O}$, and can therefore be prevented costlessly. This observation makes precise the sense in which a seller learning about the correlation structure relaxes the dynamic mechanism design setup. Moreover, it clarifies that in the two models, different incentive compatibility considerations are responsible for the economic distortions. Consequently, the type of distortions in these models differ. This lead to a failure of the generalized no distortion at the top principle.

Individual Rationality. Concerning the buyer's acceptance of the mechanism, note that a direct mechanism that satisfies the incentive constraint (5) yields a buyer of ex ante type $\theta = h$ at least his outside option of zero if and only if

$$U_h \equiv \pi[q_{hp}h + \delta Q_{hph}h - T_{hph}] + (1-\pi)[q_{hs}h + \delta Q_{hsl}l - T_{hsl}] \geq 0. \tag{7}$$

Likewise, a direct mechanism that satisfies the incentive constraint (6) yields a buyer of ex ante type $\theta = l$ at least his outside option of zero if and only if

$$U_l \equiv \pi[q_{lp}l + \delta Q_{lpl}l - T_{lpl}] + (1-\pi)[q_{ls}l + \delta Q_{lsh}h - T_{lsh}] \geq 0. \tag{8}$$

Defining a direct mechanism $\phi = (T_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}, q_{\tilde{\theta}\tilde{\gamma}}, Q_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}})$ as feasible if it satisfies the incentive compatible constraints (5) and (6) and the individual rational constraints (7) and (8), the following lemma formalizes the previous reasoning:

Lemma 1 *Suppose $\delta > 0$. Then there is no loss in focusing on direct mechanisms $\phi = (q_{\tilde{\theta}\tilde{\gamma}}, Q_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}, T_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}})$ that are feasible provided that the optimal direct mechanism ϕ^* exhibits $q^*_{ll} = q^*_{lh}$ and $q^*_{hl} = q^*_{hh}$.*

The lemma motivates to consider the problem of maximizing the seller's payoff

$$\begin{aligned}
\Pi =\, &\nu\{\pi[T_{hph} - c(q_{hp}) - \delta C(Q_{hph})] + (1-\pi)[T_{hsl} - c(q_{hs}) - \delta C(Q_{hsl})]\} \\
&+ (1-\nu)\{\pi[T_{lpl} - c(q_{lp}) - \delta C(Q_{lpl})] + (1-\pi)[T_{lsh} - c(q_{ls}) - \delta C(Q_{lsh})]\}
\end{aligned}$$

with respect to ϕ and subject to the constraints (5), (6), (7), and (8). The lemma implies that if its solution ϕ^* does not condition q on $\tilde{\gamma}$, then the solution represents the seller's profit-maximizing contract.

The Optimization Problem. Rather than working with the quantities and transfers, it is often more convenient to reformulate the problem in terms of quantities and information rents U_h and U_l, as defined in (7) and (8). In terms of these rents, the incentive compatibility constraints (5) and (6) rewrite as

$$U_h \geq U_l + \pi[q_{lp} + \delta Q_{lpl}]\Delta + (1-\pi)[q_{ls} - \delta Q_{lsh}]\Delta; \tag{IC_h}$$

$$U_l \geq U_h - \pi[q_{hp} + \delta Q_{hph}]\Delta - (1-\pi)[q_{hs} - \delta Q_{hsl}]\Delta. \tag{IC_l}$$

while the individual rationality constraints (7) and (8) simplify to

$$U_h \geq 0; \tag{IR_h}$$

$$U_l \geq 0. \tag{IR_l}$$

Rewriting the seller's ex ante expected profit of a feasible direct mechanism Π in terms of quantities and information rents, it follows that

$$\begin{aligned}\Pi = \nu\{&\pi[(hq_{hp} - c(q_{hp})) + \delta(hQ_{hph} - C(Q_{hph}))]\\ &+ (1-\pi)[(hq_{hs} - c(q_{hs})) + \delta(lQ_{hsl} - C(Q_{hsl}))] - U_h\}\\ + (1-\nu)\{&\pi[(lq_{lp} - c(q_{lp})) + \delta(lQ_{lpl} - C(Q_{lpl}))]\\ &+ (1-\pi)[(lq_{ls} - c(q_{ls})) + \delta(hQ_{lsh} - C(Q_{lsh}))] - U_l\}.\end{aligned}$$

In the remainder, I will, with a slight abuse of notation, refer to a direct mechanism as a combination $(\{q_{\tilde{\theta}\tilde{\gamma}}, Q_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}\}, U_h, U_l)$ and study the following maximization problem:

$$\mathscr{P}: \max_{(\{q_{\tilde{\theta}\tilde{\gamma}}, Q_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}\}, U_h, U_l)} \Pi \ \text{ s.t. } (IC_h), (IC_l), (IR_h), (IR_l).$$

6 Profit-maximizing contracts

It is instructive to start the analysis with considering the implementability of the perfect price discrimination outcome of efficient quantities without leaving any consumption rents. Clearly, if this outcome is implementable, then it must be profit-maximizing, as it maximizes the overall surplus and fully allocates this surplus to the seller. Hence, I start with the question whether the following direct mechanism is feasible:

$$q_{lp} = q_{ls} = q_l^*;\ q_{hp} = q_{hs} = q_h^*;\ Q_{lpl} = Q_{hsl} = Q_l^*;\ Q_{lsh} = Q_{hph} = Q_h^*;\ U_h = U_l = 0. \tag{9}$$

While the static benchmark showed that the perfect price discrimination outcome is not implementable, the following proposition shows that, when the seller perfectly learns the correlation structure, perfect price discrimination is attainable if the degree of persistence, π, is in between two thresholds $\bar{\pi}_l^*$ and $\bar{\pi}_h^*$, where[8]

$$\bar{\pi}_l^* \equiv 1 - \frac{q_h^* + \delta Q_h^*}{\delta(Q_l^* + Q_h^*)} < 1/2; \bar{\pi}_h^* \equiv 1 - \frac{q_l^* + \delta Q_l^*}{\delta(Q_l^* + Q_h^*)} < 1.$$

Proposition 1 *The seller attains perfect price discrimination if and only if $\pi \in [\bar{\pi}_l^*, \bar{\pi}_h^*]$.*[9]

In light of classical results concerning monopolistic screening, the proposition may seem surprising. Indeed, as shown in Section 4, the outcome is not implementable with $\delta = 0$.[10] In order to understand this difference, recall that in a static framework, the mechanism (9) violates the incentive constraint of the efficient type h. The reason for this is that by claiming to be the inefficient type $\theta = l$, type $\theta = h$ can secure himself a strictly positive rent. Hence, the seller has to concede a strictly positive rent to induce type $\theta = h$ to reveal his type truthfully.

In a dynamic setup in which valuations are negatively correlated, a positive rent from understating type θ is however not guaranteed, because when type $\theta = h$ turns into an inefficient type $\Theta = l$, the buyer is forced to overconsume, as he now receives the first best level of the efficient type $\Theta = h$, which exceeds his first best level as an inefficient type $\Theta = l$. Hence, negative correlation leads to a force that dissuades a type $\theta = h$ from understating his type.

This force dominates when the negative correlation is sufficiently likely. This explains why a necessary condition for the implementability of the perfect price discrimination outcome is that the likelihood of persistent types, π, lies below the threshold level $\bar{\pi}_h^*$, because correlation is negative only when types are not persistent. On the other hand, when this likelihood is too low, the force is so strong that the seller now needs to concede an information rents to type $\theta = l$. As a result, an implementability of the perfect price discrimination outcome (9) also requires that the likelihood of persistence π is not too low; it has to exceed the threshold value $\bar{\pi}_l^*$.

I next turn to analyzing the profit-maximizing contract when a negative correlation of valuations is relatively unlikely so that, just as in a static model, the first best mechanism (9) violates the incentive constraint (IC_h). This is the case when π exceeds $\bar{\pi}_h^*$. As reflected in the second benchmark of Section 4, it is well known that the binding constraints in static monopolistic screening are the incentive constraint of the efficient type and the individual rationality constraint of the inefficient type.

The next proposition confirms that this property of monopolistic screening extends to my setup when the likelihood of persistence, π, not only exceeds the threshold $\bar{\pi}_h^*$ but also the threshold

$$\bar{\pi}^h \equiv 1 - \frac{q_l^h + \delta Q_l^h}{\delta(Q_h^h + Q_l^h)},$$

where, similarly to q_l^h, the quantities Q_l^h and Q_h^h are implicitly defined by

$$C'(Q_l^h) = (l - \varphi\Delta)^+; C'(Q_h^h) = h + \varphi\Delta.$$

The next proposition fully characterizes the profit-maximizing outcome for this case.

[8]To see $\bar{\pi}_l^* < 1/2$, note that $\bar{\pi}_l^* = 1 - \frac{q_h^* + \delta Q_h^*}{\delta(Q_l^* + Q_h^*)} < 1 - \frac{\delta Q_h^*}{\delta(Q_l^* + Q_h^*)} < 1 - \frac{\delta Q_h^*}{\delta(Q_h^* + Q_h^*)} = 1/2$.

[9]Note that $\bar{\pi}_h^* \geq 0$ implies $\bar{\pi}_l^* \leq \bar{\pi}_h^*$.

[10]For values of $\delta > 0$ close to 0, it holds $\bar{\pi}_h^* < 0$ and $\bar{\pi}_h^* < 0$, implying that perfect price discrimination is infeasible.

Proposition 2 *If $\pi \geq \bar{\pi}^h$, then $\pi \geq \bar{\pi}^*_h$ and the solution exhibits both upward and downward distortions. At the optimum, only the incentive constraint (IC_h) and the participation constraint (IR_l) bind, and quantities satisfy*

$$q_{lp} = q_{ls} = q_l^h < q_{hp} = q_{hs} = q_h^*;\ Q_{lpl} = Q_l^h < Q_{hsl} = Q_l^* < Q_{hph} = Q_h^* < Q_{lsh} = Q_h^h.$$

The proposition shows that for the case $\pi \geq \bar{\pi}^h$, the optimal quantities q_l and q_h coincide with the ones under monopolistic screening as derived in the benchmark of Proposition 0. That is, the output level for the high type $\theta = h$ is efficient, $q_h = q_h^*$, and, hence, there is no distortion at the top.

In contrast, the generalized no distortion at the top principle of Battaglini (2005) does not obtain. In particular, the quantity Q_{lsh} is distorted and actually upward: $Q_{lsh} > Q_h^*$. In order to understand this upward distortion, recall that the optimal mechanism distorts quantities in order to reduce the information rents to the ex ante high type $\theta = h$ for truthtelling. To see that an upward rather than a downward distortion of Q_{lsh} reduces these rents, suppose that Q_{lsh} would be efficient, i.e. $Q_{lsh} = Q_h^*$. This means that if the high type, $\theta = h$, claims to be low by sending a report $\tilde{\theta} = l$, he receives the quantity Q_h^* in case his type switches to a low type, $\Theta = l$. However, from the perspective of a low type $\Theta = l$, the quantity $Q_{lsh} = Q_h^*$ is inefficiently high. Consequently, reducing Q_{lsh} below Q_h^* would only make a misreport $\tilde{\theta} = l$ more attractive to type $\theta = h$. In contrast by increasing Q_{lsh} beyond Q_h^*, the lie becomes less attractive and therefore reduces the information rents which the seller has to concede to type $\theta = h$ for revealing himself truthfully.

Propositions 1 and 2 leave open the seller's profit-maximizing contract for the intermediate case $\pi \in (\bar{\pi}^*_h, \bar{\pi}^h)$. For these values of π, the perfect price discrimination outcome violates the incentive constraint (IC_h), whereas the second best as characterized in Proposition 2 violates the participation constraint (IR_h). This suggest that for these intermediate values, there are three binding constraints: (IR_l), (IR_h) and (IC_h). The next proposition confirms this suggestion.

Proposition 3 *If $\pi \in (\bar{\pi}^*_h, \bar{\pi}^h)$, then the solution exhibits both upward and downward distortions. At the optimum, the incentive constraint (IC_h) and both participation constraints (IR_l) and (IR_h) are binding, and quantities satisfy*

$$q_{lp} = q_{ls} < q_l^* < q_{hp} = q_{hs} = q_h^*; Q_{lpl} < Q_{hsl} = Q_l^* < Q_{hph} = Q_h^* < Q_{lsh}.$$

In case $\bar{\pi}^*_l < 0$, Propositions 1, 2 , and 3 cover all values $\pi \in [0,1]$. Because $\bar{\pi}^*_l < 0$ is equivalent to $\delta < q_h^*/Q_l^*$, the three propositions fully characterize the profit-maximizing contract when the future is not too important.

I next extend the analysis to settings where the future is important so that $\pi < \bar{\pi}^*_l$ for values of π close to zero. Recall that for $\pi < \bar{\pi}^*_l$ the perfect price discrimination outcome 9 violates the incentive constraint (IC_l) rather than (IC_h). Mirroring the idea behind Proposition 2 to investigate the optimal contract under the condition that the incentive constraint of the type that violates this outcome is binding together with the participation constraint of the other type, I first analyze contracts for which only the incentive constraint of the inefficient type, (IC_l), and the participation constraint of the efficient type, (IR_h) are binding.

Defining the quantities q_h^l, Q_l^l, and Q_h^l implicitly by

$$c'(q_h^l) = h + \Delta/\varphi; C'(Q_l^l) = (l - \Delta/\varphi)^+; C'(Q_h^l) = h + \Delta/\varphi;$$

the next proposition derives the optimal contract for degrees of persistence smaller than $\bar{\pi}_l^*$ and

$$\bar{\pi}^l \equiv 1 - \frac{q_h^l + \delta Q_h^l}{\delta(Q_h^l + Q_l^l)}.$$

Proposition 4 *If $\pi \leq \bar{\pi}^l$, then $\pi \leq \bar{\pi}_l^*$ and the solution exhibits both upward and downward distortions. At the optimum, only the incentive constraint (IC_l) and the participation constraint (IR_h) bind and quantities satisfy*

$$q_{lp} = q_{ls} = q_l^* < q_h^* < q_{hp} = q_{hs} = q_h^l; Q_{hsl} = Q_l^l < Q_{lpl} = Q_l^* < Q_{lsh} = Q_h^* < Q_{hph} = Q_h^l.$$

For $\pi \in (\bar{\pi}^l, \bar{\pi}_l^*)$, the mechanisms identified in Proposition 1 are infeasible since they violate (IC_l), whereas the mechanisms identified in Proposition 4 are infeasible since they violate (IR_l). Mirroring the case analyzed in Proposition 3, this suggests that for the range $\pi \in (\bar{\pi}^l, \bar{\pi}_l^*)$, optimal contracts have all three constraints (IR_l), (IC_l), (IR_h) binding. The following proposition confirms this suggestion and characterizes the profit-maximizing contract.

Proposition 5 *If $\pi \in (\bar{\pi}^l, \bar{\pi}_l^*)$, then the solution exhibits both upward and downward distortions. At the optimum, the incentive constraint (IC_l) and both participation constraints (IR_l) and (IR_h) are binding, and quantities satisfy*

$$q_{lp} = q_{ls} = q_l^* < q_h^* < q_{hp} = q_{hs}; Q_{hsl} < Q_{lpl} = Q_l^* < Q_{lsh} = Q_h^* < Q_{hph}.$$

The solution in each of the previous propositions does not condition the quantities q on the seller's report about the correlation γ. Lemma 1 therefore immediately implies following corollary.

Corollary 1 *Propositions 1 to 5 completely characterize the seller's profit-maximizing direct mechanism for all possible parameter constellations. For any $\pi \notin [\bar{\pi}_l^*, \bar{\pi}_h^*]$, perfect price discrimination is unattainable and violates the generalized no distortion at the top principle, as either Q_{lsh} or Q_{hsl} is distorted.*

From an applied, real-life perspective, the profit-maximizing direct mechanisms may however look rather unnatural. I therefore close this section with addressing the question whether there are more natural contracts that implement the optimal allocations indirectly. In particular, I show that a menu of *compensation-driven partial shipment contracts* indirectly implements the outcomes associated with the optimal direct mechanisms as characterized in the previous propositions.[11] Because one can interpret a

[11]Compensation driven partial shipment (or partial delivery) clauses are standard in contract law and common in consumer contracts. They are usually motivated for reducing the seller's liability in case of unforeseen production or delivery problems, but here allow an indirect implementation of the outcome of any profit-maximizing direct mechanism.

compensation-driven partial shipment contract as an option contract on part of the seller, this indirect implementation mirrors results in dynamic mechanism design that the allocations implemented by optimal dynamic direct mechanism can, for some dynamic environments (e.g., Courty and Li, 2000), also be implemented indirectly via a menu of option contracts.

To present this indirect implementation, I define a "compensation-driven partial shipment contract" as a triple $\phi^c = (q, \overline{Q}, \underline{Q})$ with $\overline{Q} \geq \underline{Q}$ and the following interpretation. The contract ϕ^c prescribes the shipment of the quantity q of good 1 in period 1 and the quantity $\overline{Q}$ of good 2 in period 2. It however gives the seller the option of only a partial shipment of $\underline{Q}$ units of good 2 in period 2 for a compensation to the buyer of $C(\overline{Q}) - C(\underline{Q})$. That is, ϕ^c is such that the seller is indifferent about exercising her partial shipment option because her savings in production costs exactly equals the compensation that she has to pay the buyer for exercising this option.

Noting that any profit-maximizing direct mechanism $(\{q_{\tilde{\theta}\tilde{\gamma}}, Q_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}\}, U_h, U_l)$ exhibits $q_{\tilde{\theta}p} = q_{\tilde{\theta}s}$, $Q_{lsh} \geq Q_{lpl}$ and $Q_{hph} \geq Q_{hsl}$, consider the following two compensation-driven partial shipment contracts (ϕ_l^c, ϕ_h^c) associated with a profit-maximizing direct mechanism $(\{q_{\tilde{\theta}\tilde{\gamma}}, Q_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}\}, U_h, U_l)$:

$$\phi_l^c \equiv (q_{lp}, Q_{lsh}, Q_{lpl}) \text{ and } \phi_h^c \equiv (q_{hp}, Q_{hph}, Q_{hsl}).$$

Let the price of contract ϕ_l^c equal

$$T_l^c \equiv l q_{lp} + \pi(\delta l Q_{lpl} + C(Q_{lsh}) - C(Q_{lpl})) + (1-\pi)\delta h Q_{lsh} - U_l;$$

and the price of contract ϕ_h^c equal

$$T_h^c \equiv h q_{hp} + \pi \delta h Q_{hph} + (1-\pi)(\delta l Q_{hsl} + C(Q_{hph}) - C(Q_{hsl})) - U_h.$$

Now consider the seller offering the buyer the choice between (ϕ_l^c, T_l^c) and (ϕ_h^c, T_h^c), and using the following strategy to exercise her option to reduce the shipment: When the buyer picks ϕ_l^c, the seller exercises her option of partial shipment only when she learns that types are persistent. When the buyer picks ϕ_h^c, the seller exercises her option only when she learns that types switch. Exercising her option this way is optimal because of her indifference. Given the seller's behavior, it is optimal for buyer-type $\theta = l$ to pick ϕ_l^c, yielding the utility U_l, and for buyer-type $\theta = h$ to pick ϕ_h^c, yielding the utility U_h. Hence, as claimed, the menu of compensation-driven partial shipment contracts $\{(\phi_l^c, T_l^c), (\phi_h^c, T_h^c)\}$ replicates the outcome that the profit-maximizing direct mechanism $(\{q_{\tilde{\theta}\tilde{\gamma}}, Q_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}\}, U_h, U_l)$ implements.

7 Comparative Statics

In this section, I investigate the comparative statics of the profit-maximizing mechanism and its subsequent distortions. Because the previous propositions fully characterize this mechanism for all parameters π, the comparative statics follow directly from these propositions.

Starting with considering the comparative statics in the likelihood of persistent types, π, Figure 1 illustrates the optimal quantities as a function of π. It moreover links the distortions of these optimal quantities with the specific ranges for which different combinations of constraints are binding. In par-

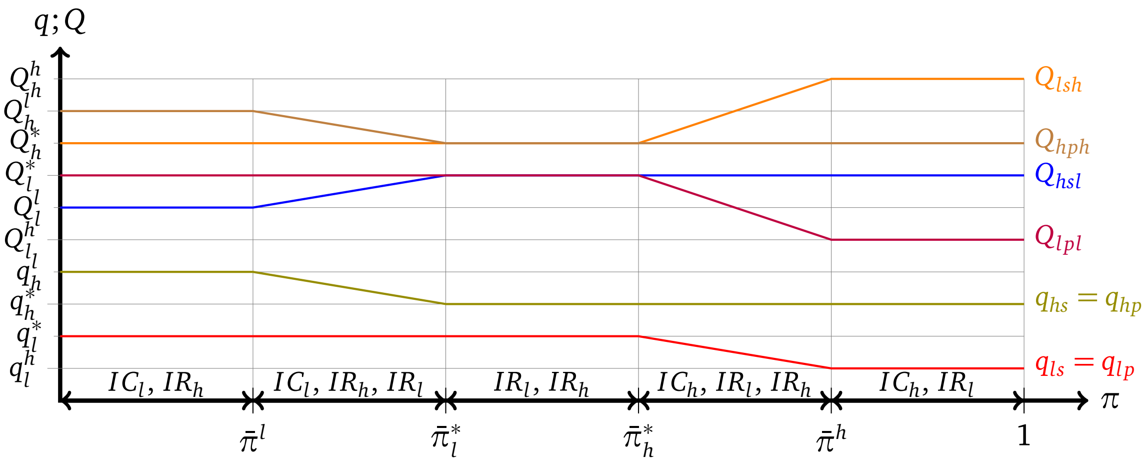

Figure 1: Optimal quantities for different degrees of persistence π

ticular, it illustrates the case when cost functions are quadratic and for which, as illustrated, the optimal quantities are piece-wise linear functions of π.

Recalling that the parameter δ measures the relative importance of good Q versus good q, I next consider the comparative statics in the parameter $\delta > 0$. Intuitively, a small δ represents a setting in which the seller's data mining techniques have little economic impact, because for δ small mainly good q matters for the payoffs. However, as δ grows large, good Q becomes more important, and the seller's learning about the correlated structure becomes the main driver of payoffs. This suggests that, as δ grows, also the seller's ability to price discriminate grows. Indeed, one may expect that, in the limit when δ grows without bounds, the seller achieves an outcome arbitrarily close to the perfect price discrimination one. An insight of this section is that, due to an informational complementarity, this intuition is misleading. By contrast, the seller attains perfect price discrimination in the limit only when, as indicated in Figure 1, the degree of persistence π lies in an intermediate range.

In order to obtain this insight, it is helpful to introduce notation that relates the thresholds $\bar{\pi}_l^*$ and $\bar{\pi}_h^*$ to δ. In particular, define the two thresholds δ_l^* and δ_h^* that correspond to the two thresholds $\bar{\pi}_l^*$ and $\bar{\pi}_h^*$:[12]

$$\delta_l^* \equiv \frac{q_h^*}{Q_l^* + \pi(Q_l^* + Q_h^*)}; \; \delta_h^* \equiv \frac{q_l^*}{Q_h^* + \pi(Q_h^* + Q_l^*)}.$$

It then follows that δ_l^* is strictly positive if and only if

$$\pi < \bar{\pi}_l^\infty \equiv \frac{Q_l^*}{Q_l^* - Q_h^*} = \lim_{\delta \to \infty} \bar{\pi}_l^*.$$

Likewise, δ_h^* is strictly positive if and only if

$$\pi < \bar{\pi}_h^\infty \equiv \frac{Q_h^*}{Q_h^* - Q_l^*} = \lim_{\delta \to \infty} \bar{\pi}_h^*.$$

[12]That is, $\pi = \bar{\pi}_i^* \Leftrightarrow \delta = \delta_i^*$.

It is straightforward to see that the limit values $\bar{\pi}_l^\infty$ and $\bar{\pi}_h^\infty$ satisfy the following ordering:

$$0 < \bar{\pi}_l^\infty < 1/2 < \bar{\pi}_h^\infty < 1.$$

Equipped with this notation, the next proposition shows that only if the degree of persistence, π, lies in between the two limit values $\bar{\pi}_l^\infty$ and $\bar{\pi}_h^\infty$, then perfect price discrimination is attainable when δ grows unbounded.

Proposition 6 *For the limit case, where δ grows unbounded, the profit-maximizing outcome coincides with perfect price discrimination if and only if $\pi \in [\bar{\pi}_l^\infty, \bar{\pi}_h^\infty]$.*

Hence, even if δ becomes unboundedly large so that the economic significance of the good q vanishes, learning the buyer's persistence does not enable the seller to extract all rents from the buyer, except for intermediate degrees of persistence. This result obtains even though, as argued in Section 5, the seller can ensure a truthful revelation of both γ and Θ at no costs in excess of the costs associated with a truthful revelation of the buyer's ex ante private information θ.

The fact that the buyer's information rents do not vanish when the good q becomes economicly insignificant depends on an informational complementarity. Indeed, from only learning whether the buyer's type is persistent or switches, the seller cannot fully deduce the value of Θ. It is only in combination with the buyer's ex ante private information that learning the persistence allows the seller to learn Θ perfectly.

As a result, the buyer's ex ante private information has two informational roles. Its first role is the usual one of restricting the seller in extracting rents concerning the consumption of good q. Its secondary role is completing the seller's information about γ and thereby enabling her to learn Θ perfectly. In line with standard intuition, the first informational role of the buyer's ex ante private information vanishes as δ becomes large. By contract, its secondary informational role gains in importance as δ increases.

To best see the complementary role of the ex ante private information, consider the case $\nu = 1/2$, where the complementarity is extreme. Note that without observing γ, the probability of $\Theta = h$ equals $\pi\nu + (1-\pi)(1-\nu)$, while the probability of $\Theta = l$ equals $\pi(1-\nu) + (1-\pi)\nu$. Given $\nu = 1/2$, both probabilities are $1/2$. Now consider the effect of the seller learning γ. If she learns $\gamma = p$, she knows that $\Theta = h$ if and only if $\theta = h$. Since the likelihood of $\theta = h$ is $\nu = 1/2$, she puts probability $1/2$ on $\Theta = h$ after observing the persistent correlation structure $\gamma = p$. Similarly, if she learns $\gamma = s$, she then knows that $\Theta = h$ if and only if $\theta = l$. But since the likelihood of $\theta = l$ is $1 - \nu = 1/2$, she also puts probability $1/2$ on $\Theta = h$ after observing $\gamma = s$. Hence, despite learning the realization of γ perfectly, her beliefs about Θ remain unchanged. This shows that a perfect signal about γ is completely uninformative about Θ. In contrast, if the seller observes both γ and θ, she is certain that $\Theta = h$ if $\gamma = p$ and $\theta = h$, or if $\gamma = s$ and $\theta = l$, while she is certain that $\Theta = l$ for the other realizations of γ and θ. Hence, perfectly learning γ is informative about Θ only if the seller also learns θ, whereas it is perfectly uninformative if the seller does not learn θ. This illustrates the informational complementary between γ and θ.

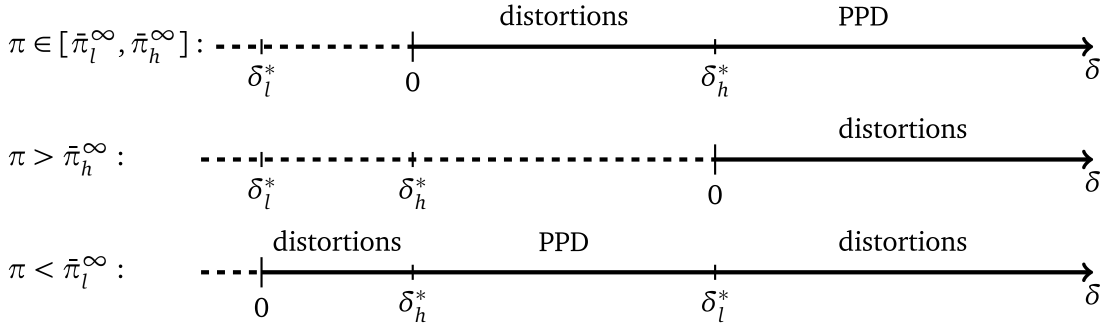

Figure 2: Attainability of perfect price discrimination (PPD) as related to δ

Intuitively, Proposition 6 follows from the fact that the thresholds $\bar{\pi}_l^*$ and $\bar{\pi}_h^*$ are increasing in δ.[13] In order to obtain further insights about the implementability of perfect price discrimination, it is instructive to consider the comparative statics in δ for the three separate cases: 1) $\pi \in [\bar{\pi}_l^{\infty}, \bar{\pi}_h^{\infty}]$ is intermediate, 2) π exceeds the threshold $\bar{\pi}_h^{\infty}$; 3) π is smaller than the threshold $\bar{\pi}_l^{\infty}$. Figure 2 illustrates these cases graphically.

The top panel of Figure 2 illustrates that in Case 1 the threshold δ_l^* is negative, while δ_h^* is positive. Hence, π exceeds $\bar{\pi}_l^*$ for all δ, but for δ small, π also exceeds $\bar{\pi}_h^*$. Hence, Proposition 3 or 2 applies, implying that the seller cannot attain perfect price discrimination. In contrast, a δ that exceeds the threshold δ_h^* implies that π is smaller than $\bar{\pi}_h^*$ but larger than $\bar{\pi}_l^*$. Proposition 1 therefore applies, demonstrating that perfect price discrimination is implementable. Perfect price discrimination obtains therefore not only in the limit but for any value of δ exceeding δ_h^*.

The second panel of Figure 2 depicts that in Case 2 both δ_l^* and δ_h^* are negative. Hence, π exceeds both $\bar{\pi}_l^*$ and $\bar{\pi}_h^*$ for any δ. For this case either Proposition 2 or 3 applies, implying that the outcome of perfect price discrimination is unattainable for any δ and also not attainable in the limit.

Finally, the lower panel of Figure 2 illustrates that in Case 3, both δ_l^* and δ_h^* are positive and, moreover, δ_l^* exceeds δ_h^*. Hence, for δ smaller than δ_h^*, it follows that π is larger than $\bar{\pi}_h^*$, implying that the outcome of perfect price discrimination is unattainable, as either Proposition 2 or 3 applies. For δ in between δ_h^* and δ_l^*, it follows that π lies in between $\bar{\pi}_l^*$ and $\bar{\pi}_h^*$, implying that Proposition 1 applies so that the outcome of perfect price discrimination is attainable. Yet, for δ exceeding δ_l^*, both $\bar{\pi}_h^*$ and $\bar{\pi}_l^*$ exceed π, implying that the outcome of perfect price discrimination is unattainable, since either Proposition 4 or 5 applies. In this final case, the comparative statics are non-monotonic in δ: perfect price discrimination is attainable for intermediate values of δ but not for δ small or δ large.

Given the full characterization of the profit-maximizing contract by Propositions 1 to 5, it is also straightforward to deduce the comparative statics in the parameter ν. Recall that the parameter ν affects the strength of the complementarity effect. For this reason, the comparative statics of ν provides insights about how this complementarity effect impacts distortions. The next lemma shows however that the complementarity effect not only depends on the parameter ν but also on the parameter π, the likelihood that the buyer has a persistent valuation.

[13] This is most easily seen after rewriting $\bar{\pi}_l^*$ as $1-\frac{q_h^*/\delta+Q_h^*}{Q_l^*+Q_h^*}$ and $\bar{\pi}_h^*$ as $1-\frac{q_l^*/\delta+Q_l^*}{Q_l^*+Q_h^*}$.

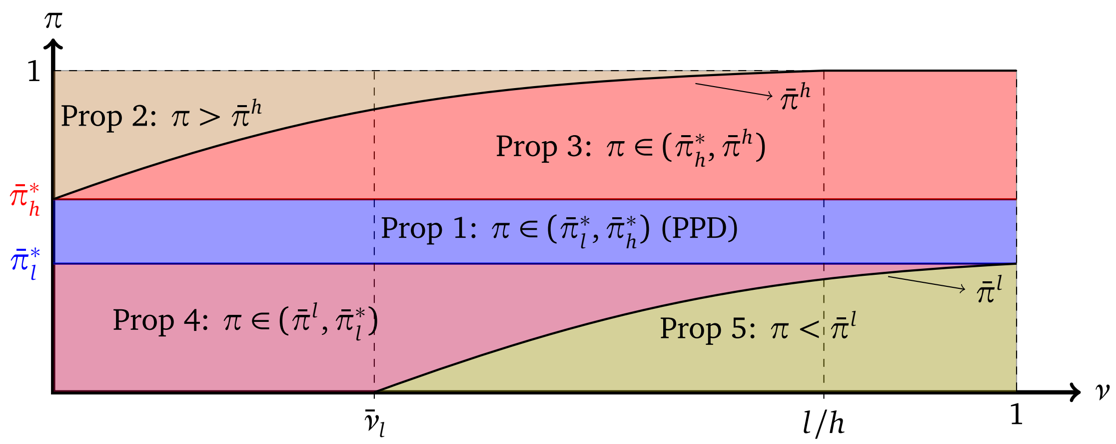

Figure 3: Comparative statics in v and π of the attainability of PPD and implied economic distortions.

Lemma 2 *The thresholds $\bar{\pi}_\ell^*$ and $\bar{\pi}_h^*$ are independent of v. The threshold $\bar{\pi}^h$ is increasing in v whenever $\bar{\pi}^h > 0$, it equals $\bar{\pi}_h^*$ for $v = 0$, and equals 1 for $v \geq l/h$. The threshold $\bar{\pi}^l$ is increasing in v whenever $\bar{\pi}^l > 0$, it converges to $\bar{\pi}_l^*$ when v approaches 1, and is smaller than 0 for $v < \bar{v}_l$, where $\bar{v}_l \in (0,1)$.*

Figure 3 illustrates the result for the case that $\bar{\pi}_l^*, \bar{\pi}_h^* > 0$. In addition, it shows how the different cases and their associated distortions, as covered in the Propositions 1 to 5, depend on the different valuations of v and π.

8 Imperfect Learning

Focusing on the implementability of perfect price discrimination, I analyzed the seller's best case that she learns the correlation structure perfectly. Assuming binary types, this means that the seller learns whether types are persistent or switch. Hence, in terms of Markov chains, the buyer's transition matrix of his valuation has a degenerated form. If types are persistent, the transition matrix equals $\mathbb{M}_p$ and, in the case types switch, it equals $\mathbb{M}_s$, where

$$\mathbb{M}_p \equiv \begin{pmatrix} \mathbb{P}\{\Theta = h|\theta = h\} & \mathbb{P}\{\Theta = l|\theta = h\} \\ \mathbb{P}\{\Theta = h|\theta = l\} & \mathbb{P}\{\Theta = l|\theta = l\} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } \mathbb{M}_s \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

From an ex ante perspective, where players only know that the transition matrix $\mathbb{M}_p$ occurs with probability π and $\mathbb{M}_s$ occurs with probability $1-\pi$, the ex ante transition matrix reflects the compound lottery

$$\mathbb{M} = \pi\mathbb{M}_p + (1-\pi)\mathbb{M}_s = \begin{pmatrix} \pi & 1-\pi \\ 1-\pi & \pi \end{pmatrix}.$$

This observation suggests that a more general model of a correlation-savvy seller obtains by considering a seller who, starting from some compounded transition matrix, privately learns additional information about the "true" transition matrix $\mathbb{M}$.

The paper's analysis of studying only the degenerated transition matrices $\mathbb{M}_p$ and $\mathbb{M}_s$ looks rather special and raises questions about the robustness of the results. For instance, Section 5 presents the key analytical step that learning the correlation relaxes incentive compatibility considerably. The optimal direct mechanism has to obey only 2 of the overall 14 dynamic incentive constraints. Moreover, these two constraints differ from the binding ones when the seller does not learn about correlations.

In this section, I argue that, by using ideas from the mechanism design literature with correlated types, the same reduction of incentive constraints obtains if the seller learns the correlation structure only imperfectly. The applicability of these ideas follow because in a model with a correlation-savvy seller, the seller's private information γ and the buyer's private information Θ are necessarily stochastically correlated (but not perfectly as in the original setup). In particular, the (imperfect) correlation allows to elicit the private information without any additional costs by the use of additional lotteries.

To illustrate this for a concrete example, suppose that for some $\pi^a \neq \pi^b$, the buyer's type transitions either via the transition matrix

$$\mathbb{M}_a = \begin{pmatrix} \pi^a & 1-\pi^a \\ 1-\pi^a & \pi^a \end{pmatrix} \text{ or } \mathbb{M}_b = \begin{pmatrix} \pi^b & 1-\pi^b \\ 1-\pi^b & \pi^b \end{pmatrix}.$$

This implies that the ex ante transition matrix is the composite

$$\mathbb{M} = \pi\mathbb{M}_a + (1-\pi)\mathbb{M}_b = \begin{pmatrix} \pi\pi^a + (1-\pi)\pi^b & \pi(1-\pi^a) + (1-\pi)(1-\pi^b) \\ \pi(1-\pi^a) + (1-\pi)(1-\pi^b) & \pi\pi^a + (1-\pi)\pi^b \end{pmatrix}.$$

Maintaining the assumption that the draws of the initial valuation $\theta \in \{l, h\}$ and the correlation structure $\gamma \in \{a, b\}$ are stochastically independent, the joint probability distribution of (θ, Θ) is

$$\mathbb{P}\{(\theta,\Theta) = (h,h)\} = \nu(\pi\pi^a + (1-\pi)\pi^b); \qquad \mathbb{P}\{(\theta,\Theta) = (h,l)\} = \nu(\pi(1-\pi^a) + (1-\pi)(1-\pi^b));$$
$$\mathbb{P}\{(\theta,\Theta) = (l,l)\} = (1-\nu)(\pi\pi^a + (1-\pi)\pi^b); \quad \mathbb{P}\{(\theta,\Theta) = (l,h)\} = (1-\nu)(\pi(1-\pi^a) + (1-\pi)(1-\pi^b)),$$

where $\mathbb{P}\{\theta = h\} = \nu$ represents once more the probability that the initial valuation θ is h rather than l.

In order to see that the seller's private information γ and the buyer's private information Θ are correlated, note that the seller believes that Θ equals θ with probability π^a after observing the correlation structure $\gamma = a$, whereas she believes this with probability π^b after observing $\gamma = b$. Because $\pi^a \neq \pi^b$, it follows that γ and Θ are stochastically correlated. Exploiting this correlation, I next show how lotteries in the spirit of Crémer and McLean (1985) allow to extend fully the analysis of the Section 5 to a seller who learns the correlation structure only imperfectly. This however requires a risk-neutral third party who is willing to make and send transfers to the buyer and the seller that are zero in expectation. In the context of a digital markets, this role can taken on by a platform.

I thereby first extend the insight of Section 5 that direct mechanisms can elicit the seller's private signal γ costlessly. For this, consider the expected payments of an additional conditional transfer schedule $\{\bar{\tau}^s_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}\} \in \mathbb{R}^8$ to the seller. That is, conditional on reports $\tilde{\theta}$, $\tilde{\gamma}$ and $\tilde{\Theta}$, the seller receives the additional payment $\bar{\tau}^s_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}$. Given that the buyer initially reported $\tilde{\theta}$ rather than $\theta' \equiv \{h, l\}\backslash\tilde{\theta}$, and the seller observes $\gamma = a$, the seller expects payment $\pi^a\tau^s_{\tilde{\theta}a\tilde{\theta}} + (1-\pi^a)\tau^s_{\tilde{\theta}a\theta'}$ from reporting $\tilde{\gamma} = a$ honestly, whereas she expects the payment $\pi^a\tau^s_{\tilde{\theta}b\tilde{\theta}} + (1-\pi^a)\tau^s_{\tilde{\theta}b\theta'}$ from misreporting $\tilde{\gamma} = b$. Similarly, after a report $\tilde{\theta}$ and observing $\gamma = b$, she expects the payment $\pi^b\tau^s_{\tilde{\theta}b\tilde{\theta}} + (1-\pi^b)\tau^s_{\tilde{\theta}b\theta'}$ from reporting $\gamma = b$ honestly and the payment $\pi^b\tau^s_{\tilde{\theta}a\tilde{\theta}} + (1-\pi^b)\tau^s_{\tilde{\theta}a\theta'}$ from misreporting $\tilde{\gamma} = a$. Because $\pi^a \neq \pi^b$, it follows that, for any K, there is a payment schedule $\{\bar{\tau}^s_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}\} \in \mathbb{R}^8$ such that the following 8 equalities are satisfied

$$\pi^a \tau^s_{\tilde{\theta}a\tilde{\theta}} + (1-\pi^a)\tau^s_{\tilde{\theta}a\theta'} = \pi^b \tau^s_{\tilde{\theta}b\tilde{\theta}} + (1-\pi^b)\tau^s_{\tilde{\theta}b\theta'} = 0,$$

and

$$\pi^a \tau^s_{\tilde{\theta}b\tilde{\theta}} + (1-\pi^a)\tau^s_{\tilde{\theta}b\theta'} = \pi^b \tau^s_{\tilde{\theta}a\tilde{\theta}} + (1-\pi^b)\tau^s_{\tilde{\theta}a\theta'} = K.$$

Setting $K < 0$ low enough ensures that the seller reports honestly her signal γ.

Next, I show that, because also the buyer's belief about the seller's private information γ depends non-trivially on his private information (θ, Θ), the insight of Section 5 concerning the buyer's relevant truthtelling constraints extends as well. To see this, first note that, for a persistent buyer-type, the buyer's Bayes' consistent belief that he is facing a seller with private information $\gamma = a$ equals

$$\alpha_p \equiv \mathbb{P}\{\gamma = a | \theta = \Theta\} = \frac{\mathbb{P}\{(\theta, \gamma, \Theta) = (\theta, a, \theta)\}}{\mathbb{P}\{(\theta, \Theta) = (\theta, \theta)\}} = \frac{\pi\pi^a}{\pi\pi^a + (1-\pi)\pi^b},$$

whereas for a switching buyer-type, it equals

$$\alpha_s \equiv \mathbb{P}\{\gamma = a | \theta \neq \Theta\} = \frac{\pi(1-\pi^a)}{\pi(1-\pi^a) + (1-\pi)(1-\pi^b)}.$$

Because $\pi^a \neq \pi^b$ implies $\alpha_p \neq \alpha_s$, it follows that the buyer's belief also depends non-trivially on his private information, expressing that types are correlated.

Exploiting this correlation, consider the schedule $\{\tau^b_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}\} \in \mathbb{R}^8$ as representing the additional transfer which the buyer has to make conditional on reports $\tilde{\theta}$, $\tilde{\gamma}$, and $\tilde{\Theta}$. Hence, after having honestly reported $\tilde{\theta} = \theta$, a persistent buyer expects to make an additional payment $\alpha_p \tau^b_{\theta a\theta} + (1-\alpha_p)\tau^b_{\theta b\theta}$ from also honestly reporting $\tilde{\Theta} = \theta$, while expecting this extra payment to be $\alpha_p \tau^b_{\theta a\theta'} + (1-\alpha_p)\tau^b_{\theta b\theta'}$ from misreporting $\tilde{\Theta} = \theta' = \{h, l\}\backslash\{\theta\}$. Similarly, after honestly reporting $\tilde{\theta} = \theta$, a switching buyer type $\Theta \neq \theta$ expects to receive the payment $\alpha_s \tau^b_{\theta a\Theta} + (1-\alpha_s)\tau^b_{\theta b\Theta}$, from honestly reporting $\tilde{\Theta} = \Theta \neq \theta$, while expecting a payment $\alpha_s \tau^b_{\theta a\theta} + (1-\alpha_s)\tau^b_{\theta b\theta}$ from misreporting $\tilde{\Theta} = \theta$. Given $\alpha_p \neq \alpha_s$, there is a transfer schedule $\{\bar{\tau}^b_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}\} \in \mathbb{R}^8$ such that, given the honest report $\tilde{\theta} = \theta$, the expected payment from reporting type Θ honestly equals 0:

$$\alpha_p \tau^b_{\theta a\theta} + (1-\alpha_p)\tau^b_{\theta b\theta} = \alpha_s \tau^b_{\theta a\theta'} + (1-\alpha_s)\tau^b_{\theta b\theta'} = 0,$$

while the buyer expects a negative transfer $K < 0$ from misreporting type Θ:

$$\alpha_p \tau^b_{\theta a\theta'} + (1-\alpha_p)\tau^b_{\theta b\theta'} = \alpha_s \tau^b_{\theta a\theta} + (1-\alpha_s)\tau^b_{\theta b\theta} = K,$$

where $\theta' = \{l, h\}\backslash\theta$. Setting K large enough acts as a punishment, ensuring that both the persistent and switching type report Θ honestly, given an honest report of θ.

The resulting transfer schedule $\{\bar{\tau}^b_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}\}$ implies that, for avoiding the punishment K, the buyer must either report θ and Θ both honestly, or misreport them both. Consequently, the lotteries associated with the transfer schedules τ^s and τ^b extend the idea of the punishments associated with the out-of-equilibrium events $\mathcal{O}$ in Section 5. Hence, also when learning is imperfect, mechanisms can costlessly prevent the seller from misreporting γ and costlessly prevent the buyer from misreporting dynamically in 12 of the 14 cases as represented by the inequalities (3) and (4). It follows that the only relevant incentive compatibility constraints of the direct mechanism $\phi = (T_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}, q_{\tilde{\theta}\tilde{\gamma}}, Q_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}})$ remain the ones that dissuade the buyer from misreporting both his initial valuation θ and his second valuation Θ.

9 Conclusion

Models of monopolistic screening yield the insight that a buyer's ex ante private information represents a countervailing power against the monopolist's ability to extract all gains from trade, guaranteeing buyers a positive information rent. Due to informational complementarities, these rents persist if the buyer faces a correlation-savvy monopolist who perfectly learns about the correlation of his valuations, and the degree of persistence is either large or small. The inability of the monopolist to achieve perfect price discrimination persists even if the economic significance of the consumption value of the good about which the consumer possesses ex ante private information vanishes.

Because this is not the case for a monopolist who learns about the buyer's valuation directly, the more general insight is that sellers who learn about correlations differ from sellers who learn about a buyer's private information. This insight cautions against a naive extrapolation from the one context to the other. Hence, in the context of big data, it matters whether retailers use data analysis for identifying robust correlations, or for learning about the private information of specific customers. Whether big data analysis is used for the former or the latter, is an empirical question, and one that is also crucial for determining the optimal regulatory response to data mining activities.

Given the stylized nature of my model, it is useful to discuss the role of its underlying assumptions in arriving at the specific results. In particular, I conclude with discussing the observability of the correlation signal, its specific timing, the role of binary types, and a possible indirect implementation of the optimal direct mechanisms.

Observability. Concerning the information structure, the setup has a monopolist who *privately* observes the correlation of valuations. An insight of the analysis is that, due to informational complementarities, this private information is inconsequential. That is, equivalent results obtain in a setup in which the correlation of valuation is revealed publicly (for instance by some third party such as a platform in the context of internet sales). The robustness of the results to the observability of the correlation signal γ obtains because, conditional on the buyer's initial information, the signal γ correlates with the buyer's signal about his latter valuation. As a result, optimal contracts can exploit the signal γ costlessly even when the seller privately observes it. Section 8 shows that in the case that the seller learns the persistence imperfectly, a costless elicitation of γ requires the use of payment lotteries as in Crémer and McLean (1985). Section 5 shows that such lotteries are not needed if the seller learns the persistence perfectly. In either case, the model is *as if* the observed correlation is verifiable and the monopolist can directly condition her contract on this information. This also implies that the monopolist does neither gain nor lose from learning the persistence in private; she has no incentive to hide this information.

Timing. I next address the timing assumption that, in the analyzed setup, the seller learns about the correlation structure *after* she offers the contract but before the buyer consumes the first good q. First, note that the fact that the optimal contract does not condition the quantity q on the seller's signal γ implies that identical results obtain if the seller learns about the correlation structure after the buyer consumes the first good q. Second, note that if the seller's learning takes place before she offers a contract, then a signalling game ensues, leading to the well-known problem of multiple equilibrium outcomes. However, the previous robustness result that identical results obtain when assuming that this correlation is revealed publicly implies both that our equilibrium outcome remains an equilibrium outcome in the signalling game, and also that it yields the seller the highest profit. Moreover, given that the profit-maximizing contract does not exploit the possibility that the initial quantity q depends on the correlation structure shows that this outcome is also attainable if the seller observes the correlation structure only after the buyer consumes q.

Binary Types. Binary valuations help to keep the model tractable and enable a full characterization of the optimal contract for all parameter constellations. This is so because with binary valuations, the correlation of types boils down to the question whether the valuation switches or remains the same. A more elaborate model with more than two types is richer, because in the case of switching it is unclear to which of the available other values the valuation switches. In other words, the dimensionality of the transition matrix increases with the number of types, which increases complexity. The binary setup is however sufficient for revealing the main driving forces of a correlation-savvy monopolist to attain first degree price discrimination: the effect that learning about whether types improve or worsen alleviates incentive constraints. This qualitative feature is best captured in a binary setup but also holds in a setup with more than two types. In the richer setup, it of course also matters how much a type worsens or improves and it is natural to expect that these magnitudes also play a role in determining the exact conditions under which first degree price discrimination is attainable.

Incomplete Contracting. As shown at the end of Section 6, the seller's profit maximizing outcome is indirectly implementable via compensation-driven partial shipment contracts. These are long term contracts and therefore require long term commitment. Clearly, the seller in the motivating beer and diaper example did not use these contracts but instead sold the two goods separately for different linear prices. Such contracts are unable to implement the identified profit-maximizing outcomes. It is an interesting, open question what type of contracts are optimal if the seller is restricted to only independent linear prices or to other forms of incomplete contracting. It is also an important and interesting question how competitive forces lead to such incomplete contracting and thereby affect the incentives of correlation-savvy sellers to engage in big data analysis. As these question lie out of the paper's scope, I leave them for future research.

Appendix

This appendix collects the proofs of the propositions and lemmas in the main text.

Proof of Proposition 0: By the revelation principle, the profit-maximizing contract is an incentive compatible mechanism $\{(q_l, p_l), (q_h, p_h)\}$, which is individual rational and maximizes the monopolist's expected profits Π. That is, it is a combination (q_l, p_l, q_h, p_h) that maximizes Π subject to (1) and (2). Considering the relaxed problem with only (IC_h) and (IR_l) yields that, for this relaxed problem, both constraints are binding, implying $p_l = lq_l$ and $p_h = hq_h - q_l\Delta$. After a substitution of these prices in the objective function, it follows that $q_h = q_h^*$ and $q_l = q_l^h$ are maximizers. It is straightforward to check that the implied direct mechanism satisfies the neglected constraints (IC_l) and (IR_h). Q.E.D.

Proof of Lemma 1: I argue that for any direct mechanisms $\phi = (q_{\tilde{\theta}\tilde{\gamma}}, Q_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}}, T_{\tilde{\theta}\tilde{\gamma}\tilde{\Theta}})$ that is feasible, there is a combination $(Q_{hsh}, Q_{lsl}, Q_{hpl}, Q_{lph}, T_{lsl}, T_{lph}, T_{hsh}, T_{hpl})$ so that the adapted direct mechanism is payoff equivalent to the original one, is feasible, and it is Bayesian incentive compatible for the seller to report γ truthfully and the buyer to report Θ truthfully. Payoff-equivalence of the adaptation is trivially ensured since given truthful reporting, the combination $(Q_{hsh}, Q_{lsl}, Q_{hpl}, Q_{lph}, T_{lsl}, T_{lph}, T_{hsh}, T_{hpl})$ does not affect the buyer's or the seller's payoff. Because of this, feasibility also trivially follows since the combination does also not affect constraints (IC_h), (IC_l), (IR_h) and (IR_l).

It remains to show that, as claimed in the body text, there is a combination $(Q_{hsh}, Q_{lsl}, Q_{hpl}, Q_{lph}, T_{lsl}, T_{lph}, T_{hsh}, T_{hpl})$ such that it is optimal for the seller to report γ truthfully if she believes that the buyer reports honestly. In the case $\gamma = p$, this requires

$$\nu\Pi(T_{hph}, q_{hp}, Q_{hph}) + (1-\nu)\Pi(T_{lpl}, q_{lp}, Q_{lpl}) \geq \nu\Pi(T_{hsh}, q_{hs}, Q_{hsh}) + (1-\nu)\Pi(T_{lsl}, q_{ls}, Q_{lsl});$$

and in the case $\gamma = s$, this requires

$$\nu\Pi(T_{hsl}, q_{hs}, Q_{hsl}) + (1-\nu)\Pi(T_{lsh}, q_{ls}, Q_{lsh}) \geq \nu\Pi(T_{hpl}, q_{hp}, Q_{hpl}) + (1-\nu)\Pi(T_{lpl}, q_{lp}, Q_{lpl}).$$

At the same time, it must be optimal for the buyer of ex ante type θ to report Θ truthfully if he reported θ honestly and believes that the seller reports the correlation structure γ honestly. For ex ante type $\theta = h$, this is the case if the following three inequalities are met

$$\begin{aligned}
&\pi U(T_{hph}, q_{hp}, Q_{hph}|h,h) + (1-\pi)U(T_{hsl}, q_{hs}, Q_{hsl}|h,l) \geq \\
&\qquad \pi U(T_{hpl}, q_{hp}, Q_{hpl}|h,h) + (1-\pi)U(T_{hsl}, q_{hs}, Q_{hsl}|h,l); \\
&\pi U(T_{hph}, q_{hp}, Q_{hph}|h,h) + (1-\pi)U(T_{hsl}, q_{hs}, Q_{hsl}|h,l) \geq \\
&\qquad \pi U(T_{hph}, q_{hp}, Q_{hph}|h,h) + (1-\pi)U(T_{hsh}, q_{hs}, Q_{hsh}|h,l); \\
&\pi U(T_{hph}, q_{hp}, Q_{hph}|h,h) + (1-\pi)U(T_{hsl}, q_{hs}, Q_{hsl}|h,l) \geq \\
&\qquad \pi U(T_{hpl}, q_{hp}, Q_{hpl}|h,h) + (1-\pi)U(T_{hsh}, q_{hs}, Q_{hsh}|h,l).
\end{aligned}$$

For ex ante type $\theta = l$, this is the case if the following three inequalities are met

$$\begin{aligned} &\pi U(T_{lpl}, q_{lp}, Q_{lpl}|l,l) + (1-\pi)U(T_{lsh}, q_{ls}, Q_{hsl}|l,h) \geq \\ &\qquad \pi U(T_{lph}, q_{lp}, Q_{lph}|l,l) + (1-\pi)U(T_{lsh}, q_{ls}, Q_{lsh}|l,h); \\ &\pi U(T_{lpl}, q_{lp}, Q_{lpl}|l,l) + (1-\pi)U(T_{lsh}, q_{ls}, Q_{hsl}|l,h) \geq \\ &\qquad \pi U(T_{lpl}, q_{lp}, Q_{lpl}|l,l) + (1-\pi)U(T_{lsl}, q_{ls}, Q_{lsl}|l,h); \\ &\pi U(T_{lpl}, q_{lp}, Q_{lpl}|l,l) + (1-\pi)U(T_{lsh}, q_{ls}, Q_{hsl}|l,h) \geq \\ &\qquad \pi U(T_{lph}, q_{lp}, Q_{lph}|l,l) + (1-\pi)U(T_{lsl}, q_{ls}, Q_{lsl}|l,h). \end{aligned}$$

Each of these 8 inequalities is implied if the following 8 conditions on the state-by-state utilities hold simultaneously

$$\begin{gathered} \Pi(T_{hph}, q_{hp}, Q_{hph}) \geq \Pi(T_{hsh}, q_{hs}, Q_{hsh}) \text{ and } \Pi(T_{lpl}, q_{lp}, Q_{lpl}) \geq \Pi(T_{lsl}, q_{ls}, Q_{lsl}); \\ \Pi(T_{hsl}, q_{hs}, Q_{hsl}) \geq \Pi(T_{hpl}, q_{hp}, Q_{hpl}) \text{ and } \Pi(T_{lsh}, q_{ls}, Q_{lsh}) \geq \Pi(T_{lph}, q_{lp}, Q_{lph}); \\ U(T_{hph}, q_{hp}, Q_{hph}|h,h) \geq U(T_{hpl}, q_{hp}, Q_{hpl}|h,h) \text{ and } U(T_{hsl}, q_{hs}, Q_{hsl}|h,l) \geq U(T_{hsh}, q_{hs}, Q_{hsh}|h,l); \\ U(T_{lpl}, q_{lp}, Q_{lpl}|l,l) \geq U(T_{lph}, q_{lp}, Q_{lph}|l,l) \text{ and } U(T_{lsh}, q_{ls}, Q_{hsl}|l,h) \geq U(T_{lsl}, q_{ls}, Q_{lsl}|l,h). \end{gathered}$$

Hence, it suffices to show that for any 8 numbers $(K_1, \ldots, K_8)$, there is a combination $(Q_{hsh}, Q_{lsl}, Q_{hpl}, Q_{lph}, T_{lsl}, T_{lph}, T_{hsh}, T_{hpl})$ such that

$$\begin{gathered} K_1 \geq \Pi(T_{hsh}, q_{hs}, Q_{hsh}) \text{ and } K_2 \geq \Pi(T_{lsl}, q_{ls}, Q_{lsl}) \\ K_3 \geq \Pi(T_{hpl}, q_{hp}, Q_{hpl}) \text{ and } K_4 \geq \Pi(T_{lph}, q_{lp}, Q_{lph}); \\ K_5 \geq U(T_{hpl}, q_{hp}, Q_{hpl}|h,h) \text{ and } K_6 \geq U(T_{hsh}, q_{hs}, Q_{hsh}|h,l). \\ K_7 \geq U(T_{lph}, q_{lp}, Q_{lph}|l,l) \text{ and } K_8 \geq U(T_{lsl}, q_{ls}, Q_{lsl}|l,h). \end{gathered}$$

To show this is indeed the case, first regroup these 8 inequalities as follows:

$$\begin{gathered} K_1 \geq \Pi(T_{hsh}, q_{hs}, Q_{hsh}) \text{ and } K_6 \geq U(T_{hsh}, q_{hs}, Q_{hsh}|h,l); \\ K_2 \geq \Pi(T_{lsl}, q_{ls}, Q_{lsl}) \text{ and } K_8 \geq U(T_{lsl}, q_{ls}, Q_{lsl}|l,h); \\ K_3 \geq \Pi(T_{hpl}, q_{hp}, Q_{hpl}) \text{ and } K_5 \geq U(T_{hpl}, q_{hp}, Q_{hpl}|h,h); \\ K_4 \geq \Pi(T_{lpl}, q_{lp}, Q_{lpl}) \text{ and } K_7 \geq U(T_{lph}, q_{lp}, Q_{lph}|l,l). \end{gathered}$$

so that taking each pair of inequalities together implies the following four necessary conditions:

$$\begin{aligned} K_1 + K_6 &\geq \Pi(T_{hsh}, q_{hs}, Q_{hsh}) + U(T_{hsh}, q_{hs}, Q_{hsh}|h,l) = \theta q_{hs} - c(q_{hs}) + \delta(\Theta Q_{hsh} - C(Q_{hsh})); \\ K_2 + K_8 &\geq \Pi(T_{lsl}, q_{ls}, Q_{lsl}) + U(T_{lsl}, q_{ls}, Q_{lsl}|l,h) = \theta q_{ls} - c(q_{ls}) + \delta(\Theta Q_{lsl} - C(Q_{lsl})); \\ K_3 + K_5 &\geq \Pi(T_{hpl}, q_{hp}, Q_{hpl}) + U(T_{hpl}, q_{hp}, Q_{hpl}|h,h) = \theta q_{hp} - c(q_{hp}) + \delta(\Theta Q_{hpl} - C(Q_{hpl})); \\ K_4 + K_7 &\geq \Pi(T_{lpl}, q_{lp}, Q_{lpl}) + U(T_{lph}, q_{lp}, Q_{lph}|l,l) = \theta q_{lp} - c(q_{lp}) + \delta(\Theta Q_{lph} - C(Q_{lph})). \end{aligned}$$

The convexity of $C(Q)$ and the assumption $\lim_{Q\to\infty} C'(Q) = \infty$ imply that for any combination $(K_1, \ldots, K_8)$ and $(q_{hs}, q_{ls}, q_{hp}, q_{lp})$ and $\delta > 0$, these four inequalities are satisfied for $(Q_{hsh}, Q_{lsl}, Q_{hpl}, Q_{lph})$ large enough. It follows that there are then also constants $(T_{hsh}, T_{lsl}, T_{hpl}, T_{lph})$ so that together with these large $(Q_{hsh}, Q_{lsl}, Q_{hpl}, Q_{lph})$ each of the previous 8 inequalities holds. Q.E.D.

Proof of Proposition 1: I verify that the efficient quantities together with $U_h = U_l = 0$ satisfy the incentive constraints IC_h and IC_l if and only if $\pi \in [\bar{\pi}_l^*, \bar{\pi}_h^*]$.

First, note that IC_h associated with this outcome simplifies to

$$\pi[q_l^* + \delta Q_l^*]\Delta + (1-\pi)[q_l^* - \delta Q_h^*]\Delta \leq 0$$

which is equivalent to $\pi \leq \bar{\pi}_h^*$.

Next, note that IC_l associated with the first best simplifies to

$$0 \geq -\pi[q_h^* + \delta Q_h^*] - (1-\pi)[q_h^* - \delta Q_l^*]$$

which is equivalent to $\pi \geq \bar{\pi}_l^*$.

The proposition follows from noting that $q_h^* > q_l^*$ and $Q_h^* > Q_l^*$ implies $\bar{\pi}_l^* < \bar{\pi}_h^*$. Q.E.D.

Proof of Proposition 2: I first prove that $\pi \geq \bar{\pi}^h$ implies $\pi \geq \bar{\pi}_h^*$. Note that, due to $\pi \geq 0$, this trivially holds when $\bar{\pi}^h < 0$ and $\bar{\pi}_h^* < 0$. So suppose $\bar{\pi}^h \geq 0$ or $\bar{\pi}_h^* \geq 0$ holds. The first inequality implies that $\delta \geq \delta^h \equiv q_l^h/Q_h^h$, whereas the second implies that $\delta \geq \delta_h^* \equiv q_l^*/Q_h^*$. Note that $\delta^h < \delta_h^*$ since $q_l^h < q_l^*$ and $Q_h^h > Q_h^*$. Because both $\bar{\pi}^h$ and $\bar{\pi}_h^*$ are increasing in δ, it follows that $\bar{\pi}_h^* \geq 0$ implies $\bar{\pi}^h \geq 0$. Hence, the statement $\bar{\pi}^h \geq 0$ or $\bar{\pi}_h^* \geq 0$ implies $\delta \geq \delta^h$.

I next show that for $\delta \geq \delta^h$, it follows $\bar{\pi}^h > \bar{\pi}_h^*$. To see this, define

$$\Delta^h(\delta) \equiv \bar{\pi}^h s \bar{\pi}_h^* = \frac{q_l^* + \delta Q_l^*}{\delta(Q_l^* + Q_h^*)} - \frac{q_l^h + \delta Q_l^h}{\delta(Q_h^h + Q_l^h)} = \frac{(q_l^*/\delta + Q_l^*)(Q_h^h + Q_l^h) - (q_l^h/\delta + Q_l^h)(Q_h^* + Q_l^*)}{(Q_h^* + Q_l^*)(Q_h^h + Q_l^h)}$$

so that it follows

$$\frac{\partial \Delta^h}{\partial \delta} = \frac{q_l^h(Q_h^* + Q_l^*) - q_l^*(Q_h^h + Q_l^h)}{\delta^2(Q_h^* + Q_l^*)(Q_h^h + Q_l^h)}.$$

Hence, the sign of $\partial \Delta^h/\partial \delta$ does not depend on δ. If the sign of $\partial \Delta^h/\partial \delta$ is positive, then recall that at $\delta = \delta^h$ it holds $\bar{\pi}^h = 0$ and $\bar{\pi}_h^* < 0$ so that $\Delta^h(\delta^h) = 0 - \bar{\pi}_h^* > 0$. The positive sign of $\partial \Delta^h/\partial \delta$ then implies $\Delta^h(\delta) > 0$ for all $\delta \geq \delta^h$ and, hence, $\bar{\pi}^h > \bar{\pi}_h^*$.

If the sign of $\partial \Delta^h/\partial \delta$ is negative, the result $\bar{\pi}^h \geq \bar{\pi}_h^*$ then follows from

$$\Delta^h(\delta) \geq \lim_{\delta \to \infty} \Delta^h(\delta) = \frac{Q_l^*(Q_h^h + Q_l^h) - Q_l^h(Q_h^* + Q_l^*)}{(Q_h^* + Q_l^*)(Q_h^h + Q_l^h)} = \frac{Q_l^* Q_h^h - Q_l^h Q_h^*}{(Q_h^* + Q_l^*)(Q_h^h + Q_l^h)} > 0,$$

where the last inequality follows from $Q_l^* > Q_l^h > 0$ and $Q_h^h > Q_h^* > 0$.

Hence, $\pi \geq \bar{\pi}^h$ implies $\pi \geq \bar{\pi}_h^*$ so that the first best violates (IC_h). I next show that $\pi \geq \bar{\pi}^h$ implies that only (IC_h) and (IR_l) bind at the maximum. Indeed, a binding incentive constraint (IC_h) and participation constraint (IR_l) imply

$$U_l = 0; U_h = \pi[q_{lp} + \delta Q_{lpl}]\Delta + (1-\pi)[q_{ls} - \delta Q_{lsh}]\Delta.$$

Substitution of these values for U_h and U_l in Π, implies maximizing

$$\begin{aligned}
&\nu\{\pi[(hq_{hp}-c(q_{hp}))+\delta(hQ_{hph}-C(Q_{hph}))]+(1-\pi)[(hq_{hs}-c(q_{hs}))+\delta(lQ_{hsl}-C(Q_{hsl}))]\}\\
&+(1-\nu)\{\pi[((l-\varphi\Delta)q_{lp}-c(q_{lp}))+\delta((l-\varphi\Delta)Q_{lpl}-C(Q_{lpl}))]\}\\
&+(1-\nu)\{(1-\pi)[((l-\varphi\Delta)q_{ls}-c(q_{ls}))+\delta((h+\varphi\Delta)Q_{lsh}-C(Q_{lsh}))]\}
\end{aligned}$$

with solution $q_{lp}=q_{ls}=q_l^h<q_{hp}=q_{hs}=q_h^*;Q_{lpl}=Q_l^h<Q_{hsl}=Q_l^*<Q_{hph}=Q_h^*<Q_{lsh}=Q_h^h$. To see that for this solution, (IC_l) holds, note that for this solution (IC_l) simplifies to

$$0\geq\pi[q_l^h+\delta Q_l^h]\Delta+(1-\pi)[q_l^h-\delta Q_h^h]\Delta-\pi[q_h^*+\delta Q_h^*]\Delta-(1-\pi)[q_h^*-\delta Q_l^*]\Delta,$$

and rewrites as

$$\pi[q_h^*-Q_l^h+\delta(Q_h^*-Q_l^h)]+(1-\pi)[q_h^*-Q_l^h+\delta(Q_h^h-Q_l^*)]\geq 0,$$

which holds since $q_h^*>q_l^h$ and $Q_h^h>Q_h^*>Q_l^*>Q_l^h$.

Moreover, the solution satisfies the neglected (IR_h) if and only if

$$U_h=\pi[q_l^h+\delta Q_l^h]\Delta+(1-\pi)[q_l^h-\delta(Q_l^h+Q_h^h-Q_l^h)]\Delta=(q_l^h+\delta Q_l^h)\Delta-(1-\pi)\delta(Q_h^h+Q_l^h)\Delta\geq 0,$$

which if equivalent to $\pi\geq\bar{\pi}^h$. Q.E.D.

Proof of Proposition 3: Assuming the three binding constraints are (IR_l), (IR_h) and (IC_h), it follows

$$U_l=0;U_h=0;0=\pi[q_{lp}+\delta Q_{lpl}]\Delta+(1-\pi)[q_{ls}-\delta Q_{lsh}]\Delta.$$

Substitution of $U_h=U_l=0$ implies to maximize

$$\begin{aligned}
W_3\equiv&\nu\{\pi[(hq_{hp}-c(q_{hp}))+\delta(hQ_{hph}-C(Q_{hph}))]+(1-\pi)[(hq_{hs}-c(q_{hs}))+\delta(lQ_{hsl}-C(Q_{hsl}))]\}\\
&+(1-\nu)\{\pi[(lq_{lp}-c(q_{lp}))+\delta(lQ_{lpl}-C(Q_{lpl}))]+(1-\pi)[(lq_{ls}-c(q_{ls}))+\delta(hQ_{lsh}-C(Q_{lsh}))]\}\\
&\text{s.t. }(1-\nu)[\pi[q_{lp}+\delta Q_{lpl}]+(1-\pi)[q_{ls}-\delta Q_{lsh}]]=0.
\end{aligned}$$

The associated Lagrangian is

$$\begin{aligned}
\mathscr{L}\equiv&\nu\{\pi[(hq_{hp}-c(q_{hp}))+\delta(hQ_{hph}-C(Q_{hph}))]+(1-\pi)[(hq_{hs}-c(q_{hs}))+\delta(lQ_{hsl}-C(Q_{hsl}))]\}\\
&+(1-\nu)\{\pi[((l-\lambda)q_{lp}-c(q_{lp}))+\delta((l-\lambda)Q_{lpl}-C(Q_{lpl}))]\\
&\qquad+(1-\pi)[((l-\lambda)q_{ls}-c(q_{ls}))+\delta((h+\lambda)Q_{lsh}-C(Q_{lsh}))]\},
\end{aligned}$$

where λ is the lagrange multiplier. Hence, the optimality conditions w.r.t. $q_{hp},Q_{hph},q_{hs},Q_{hsl}$ imply $q_{hp}=Q_{hph}=q_{hs}=q_h^*$, and $Q_{hsl}=q_l^*$; the optimality conditions w.r.t. q_{lp},Q_{lpl},q_{ls} coincide; they satisfy

$$c'(q_{lp})=C'(Q_{lpl})=c'(q_{ls})=(l-\lambda)^+;$$

where for Q_{lsh} it optimally holds

$$C'(Q_{lsh})=(h+\lambda)^+.$$

In order to see that the sign of the Lagrange multiplier λ is positive, note that with (IR_h) and (IR_l) binding (IC_h) rewrites as

$$\pi[q_{lp} + \delta Q_{lpl}]\Delta + (1-\pi)[q_{ls} - \delta Q_{lsh}]\Delta \leq 0.$$

Hence, the constraint is relaxed when the RHS rises. Consequently, the shadow cost of the constraint is positive, implying $\lambda > 0$.

Consequently, the solution exhibits

$$q_{lp} = q_{ls} < q_l^* < q_{hp} = q_{hs} = q_h^*; Q_{lpl} < Q_{hsl} = Q_l^* < Q_{hph} = Q_h^* < Q_{lsh}.$$

Finally, I check that for this solution (IC_l) is satisfied, which is the case if

$$0 \geq -\pi[q_h^* + \delta Q_h^*] - (1-\pi)[q_h^* - \delta Q_l^*]$$

which simplifies to $\pi \geq \bar{\pi}_l^*$ and holds because of $\bar{\pi}_h^* > \bar{\pi}_l^*$ and the proposition's assumption $\pi > \bar{\pi}_h^*$. Q.E.D

Proof of Proposition 4: If $\pi \leq \bar{\pi}^l$, then, necessarily, $\bar{\pi}^l \geq 0$, implying $\delta \geq \delta^l \equiv q_h^l/Q_l^l$. Note that $\pi_l^* = 0$ for $\delta = \delta_l^* \equiv q_h^*/Q_h^*$. Because $q_h^* < q_h^l$ and $Q_h^* > Q_l^l$, it follows $\delta_l^* < \delta^l$. Since π_l^* is increasing in δ, it follows that $\pi_l^* > 0$ for $\delta = \delta^l$.

I next show that for $\delta \geq \delta^l$, it holds $\bar{\pi}^l < \bar{\pi}_l^*$. To see this, define

$$\Delta^l(\delta) \equiv \bar{\pi}_l^* s \bar{\pi}^l = \frac{q_h^l + \delta Q_h^l}{\delta(Q_h^l + Q_l^l)} - \frac{q_h^* + \delta Q_h^*}{\delta(Q_l^* + Q_h^*)} = \frac{(q_h^l/\delta + Q_h^l)(Q_l^* + Q_h^*) - (q_h^*/\delta + Q_h^*)(Q_h^l + Q_l^l)}{(Q_h^l + Q_l^l)(Q_l^* + Q_h^*)}$$

so that

$$\frac{\partial \Delta^l}{\partial \delta} = \frac{q_h^l(Q_l^* + Q_h^*) - q_h^*(Q_h^l + Q_l^l)}{\delta^2(Q_l^l + Q_h^l)(Q_l^* + Q_h^*)}.$$

Hence, the sign of $\partial \Delta^l/\partial \delta$ does not depend on δ and is either positive or negative. Recall that at $\delta = \delta^l$ it holds $\bar{\pi}^l = 0$ and $\bar{\pi}_l^* > 0$ so that $\Delta^l(\delta^l) = \bar{\pi}_l^* s 0 > 0$. Hence, if the sign of $\partial \Delta^l/\partial \delta$ is positive, it follows $\Delta^l(\delta) > 0$ for all $\delta \geq \delta^l$ and, hence, $\bar{\pi}^l < \bar{\pi}_l^*$. If, on the other hand, the sign of $\partial \Delta^l/\partial \delta$ is negative, $\bar{\pi}^l < \bar{\pi}_l^*$ results from

$$\Delta^l(\delta) \geq \lim_{\delta \to \infty} \Delta^l(\delta) = \frac{Q_h^l(Q_l^* + Q_h^*) - Q_h^*(Q_h^l + Q_l^l)}{(Q_h^l + Q_l^l)(Q_l^* + Q_h^*)} = \frac{Q_h^l Q_l^* - Q_h^* Q_l^l}{(Q_h^l + Q_l^l)(Q_l^* + Q_h^*)} > 0,$$

where the last inequality follows from $Q_h^l > Q_h^* > 0$ and $Q_l^* > Q_l^l > 0$. It follows that $\pi \leq \bar{\pi}^l$ implies $\pi \leq \bar{\pi}_l^*$.

I next show that for $\pi \leq \bar{\pi}^l$, the profit-maximizing contract has (IC_l) and (IR_h) binding, while (IC_h) and (IR_l) are automatically satisfied. That is, it exhibits

$$U_h = 0; U_l = -\pi[q_{hp} + \delta Q_{hph}]\Delta - (1-\pi)[q_{hs} - \delta Q_{hsl}]\Delta$$

Indeed, substituting these values into Π leads to maximizing

$$\nu\{\pi[((h+\delta/\tilde{\nu})q_{hp}-c(q_{hp}))+\delta((h+\delta/\tilde{\nu})Q_{hph}-C(Q_{hph}))]$$
$$+(1-\pi)[((h+\delta/\tilde{\nu})q_{hs}-c(q_{hs}))+\delta((l-\delta/\tilde{\nu})Q_{hsl}-C(Q_{hsl}))]\}$$
$$+(1-\nu)\{\pi[(lq_{lp}-c(q_{lp}))+\delta(lQ_{lpl}-C(Q_{lpl}))]+(1-\pi)[(lq_{ls}-c(q_{ls}))+\delta(hQ_{lsh}-C(Q_{lsh}))]\}$$

with optimal values

$$q_{lp}=q_{ls}=q_l^*<q_h^*<q_{hp}=q_{hs}=q_h^l;Q_{hsl}=Q_l^l<Q_{lpl}=Q_l^*<Q_{lsh}=Q_h^*<Q_{hph}=Q_h^l;$$

It remains to be checked whether this solution satisfies (IC_h) and (IR_l).

To see that (IC_h) is satisfied at this solution, note that the constraint for this solution simplifies to

$$(q_h^l-Q_l^*+\delta(Q_h^l-Q_l^*))\geq(1-\pi)\delta(Q_h^l+Q_l^l-Q_l^*-Q_h^*).$$

Note that the inequality holds for any π if it holds for $\pi=0$. In this case, the inequality is equivalent to

$$(q_h^l-Q_l^*)\geq\delta(Q_l^l-Q_h^*),$$

which holds since the left hand side is positive, whereas the right hand side is negative.

Next consider (IR_l) for the solution, stating that

$$U_l=-\pi[q_h^l+\delta Q_h^l]\Delta-(1-\pi)[q_h^l-\delta Q_l^l]\Delta=[(1-\pi)\delta(Q_h^l+Q_l^l)-(q_h^l+\delta Q_h^l)]\Delta$$

is non-negative, which is a condition that rewrites as

$$\pi\leq 1-\frac{q_h^l+\delta Q_h^l}{\delta(Q_h^l+Q_l^l)}=\bar{\pi}^l.$$

Q.E.D.

Proof of Proposition 5: Assuming the three binding constraints are (IR_l), (IR_h) and (IC_l), it follows

$$U_l=0;U_h=0;\pi[q_{hp}+\delta Q_{hph}]\Delta+(1-\pi)[q_{hs}-\delta Q_{hsl}]\Delta=0$$

Substitution of $U_h=U_l=0$ implies to maximize

$$W_3 \text{ s.t. } \nu[\pi[q_{hp}+\delta Q_{hph}]+(1-\pi)[q_{hs}-\delta Q_{hsl}]]=0,$$

with W_3 as defined in the proof of Proposition 3.

The associated Lagrangian is

$$\mathcal{L}\equiv\nu\{\pi[(hq_{hp}-c(q_{hp}))+\delta(hQ_{hph}-C(Q_{hph}))]+(1-\pi)[(hq_{hs}-c(q_{hs}))+\delta(lQ_{hsl}-C(Q_{hsl}))]\}$$
$$+(1-\nu)\{\pi[(lq_{lp}-c(q_{lp}))+\delta(lQ_{lpl}-C(Q_{lpl}))]+(1-\pi)[(lq_{ls}-c(q_{ls}))+\delta(hQ_{lsh}-C(Q_{lsh}))]\}$$
$$-\lambda(\nu[\pi[q_{hp}+\delta Q_{hph}]+(1-\pi)[q_{hs}-\delta Q_{hsl}]]),$$

where λ is the lagrange multiplier. Hence, the optimality conditions w.r.t. $q_{lp}, Q_{lpl}, q_{ls}, Q_{lsh}$ imply $q_{lp} = Q_{lpl} = q_{ls} = q_l^*$, and $Q_{lsh} = q_h^*$; the optimality conditions w.r.t. q_{hp}, Q_{hph}, q_{hs} coincide; they satisfy

$$c'(q_{hp}) = c'(Q_{hph}) = c'(q_{hs}) = (h - \lambda)^+;$$

where for Q_{hsl} it optimally holds

$$c'(Q_{hsl}) = (l + \lambda)^+.$$

In order to see that the sign of the Lagrange mulitplier λ is negative, note that with (IR_h) and (IR_l) binding (IC_l) rewrites as

$$\pi[q_{hp} + \delta Q_{hph}]\Delta + (1-\pi)[q_{hs} - \delta Q_{hsl}]\Delta \geq 0$$

Hence, the constraint is strengthened when the RHS rises. Consequently, the shadow cost of the constraint is negative as raising the right hand side lowers the objective. Hence, $\lambda < 0$.

As a result, the solution exhibits

$$q_{lp} = q_{ls} = q_l^* < q_h^* < q_{hp} = q_{hs}; Q_{hsl} < Q_{lpl} = Q_l^* < Q_{lsh} = Q_h^* < Q_{hph}.$$

Finally, I check that for this solution (IC_h) is satisfied, which is the case if

$$0 \geq \pi[q_l^* + \delta Q_l^*] + (1-\pi)[q_l^* - \delta Q_h^*],$$

which simplifies to $\pi \leq \bar{\pi}_h^*$ and holds because of $\bar{\pi}_l^* < \bar{\pi}_h^*$ and the proposition's assumption $\pi < \bar{\pi}_l^*$. Q.E.D

Proof of Proposition 6: First, consider $\pi \in [0, \bar{\pi}_l^\infty)$. In this case, $\delta_l^* > 0$ so that $\pi < \pi_l^*$ for $\delta > \delta_l^*$ As a result, the optimal contract is characterized by either Proposition 4 or 5. In the first case, it holds

$$Q_{hsl} = Q_l^l < Q_{lpl} = Q_l^* < Q_{lsh} = Q_h^* < Q_{hph} = Q_h^l.$$

so that the result follows directly. In the second, case, it holds for any $\delta > \delta_l^*$ that

$$Q_{hsl} < Q_{lpl} = Q_l^* < Q_{lsh} = Q_h^* < Q_{hph}$$

so that the result follows when Q_{hsl} is decreasing and Q_{hph} is increasing in δ. To see that this is indeed the case, note that at the solution the constraint

$$\pi q_{hp} + (1-\pi)q_{hs} + \delta(\pi Q_{hph} - (1-\pi)Q_{hsl}) \geq 0$$

binds so that the Lagrange multiplier is strictly negative. Moreover, it holds $\pi Q_{hph} - (1-\pi)Q_{hsl} < 0$ so that the constraint tightens as δ grows, meaning that the Lagrange multiplier λ becomes more negative. As a result, the upward distortion on Q_{hph} and the downward distortion on Q_{hsl} intensify as δ rises. This confirms that Q_{hsl} as derived in Proposition 5 is decreasing and Q_{hph} as derived in Proposition 5 is increasing in δ and the results follows.

Second, consider $\pi \in (\bar{\pi}_l^\infty, \bar{\pi}_h^\infty]$. In this case, $\delta_l^* < 0$ and $\delta_h^* \geq 0$ so that $\pi \in (\bar{\pi}_l^\infty, \bar{\pi}_h^\infty]$ for $\delta > \delta_h^*$ As a result, the optimal contract is characterized by Proposition 1, which yields the result.

Finally, consider $\pi \in (\bar{\pi}_h^\infty, 1]$. In this case, $\delta_l^*, \delta_h^* < 0$ so that $\pi > \pi_h^*$ for $\delta > 0$ As a result, the optimal contract is characterized by either Proposition 2 or 3. In the first case, it holds

$$Q_{lpl} = Q_l^h < Q_{hsl} = Q_l^* < Q_{hph} = Q_h^* < Q_{lsh} = Q_h^h$$

so that the result follows directly. In the second case, it holds

$$Q_{lpl} < Q_{hsl} = Q_l^* < Q_{hph} = Q_h^* < Q_{lsh}$$

for any $\delta > 0$ so that the result follows when Q_{lpl} is decreasing and Q_{lsh} is increasing in δ.To see that this is indeed the case, note that at the solution the constraint

$$\pi q_{lp} + (1-\pi) q_{ls} + \delta(\pi Q_{lpl} - (1-\pi) Q_{lsh}) \leq 0$$

binds. Hence, the associated Lagrange multiplier is strictly positive and, moreover, it holds $\pi Q_{lpl} - (1-\pi) Q_{lsh} < 0$. The constraint therefore tightens as δ grows, meaning that the Lagrange multiplier λ becomes more positive. Consequently, the upward distortion on Q_{lsh} and the downward distortion on Q_{lpl} intensify as δ rises. This confirms that Q_{lpl} as derived in Proposition 3 is decreasing and Q_{lsh} as derived in Proposition 3 is increasing in δ so that the result follows. Q.E.D.

Proof of Lemma 2: The first observation follows directly from the definition of $\bar{\pi}_\ell^*$ and $\bar{\pi}_h^*$. Moreover, using the implicit definitions of q_l^h, Q_l^h, it follows that these quantities are positive if and only if $\nu < l/h$. Hence, $\bar{\pi}^h < 1$ if and only if $\nu < l/h$. For $\nu = 0$, it follows that q_l^h, Q_l^h, and Q_h^h are efficient so that $\bar{\pi}^h = \bar{\pi}_h^*$. Furthermore, it follows

$$\frac{d\bar{\pi}^h}{d\nu} = \underbrace{\frac{\partial \bar{\pi}^h}{\partial q_l^h}}_{\leq 0} \underbrace{\frac{\partial q_l^h}{\partial \nu}}_{\leq 0} + \underbrace{\frac{\partial \bar{\pi}^h}{\partial Q_l^h}}_{\leq 0} \underbrace{\frac{\partial Q_l^h}{\partial \nu}}_{\leq 0} + \underbrace{\frac{\partial \bar{\pi}^h}{\partial Q_h^h}}_{\geq 0} \underbrace{\frac{\partial Q_h^h}{\partial \nu}}_{\geq 0} \geq 0,$$

where $\partial \bar{\pi}^h / \partial Q_l^h \leq 0$ follows from $\bar{\pi}^h > 0$.

Likewise, using the implicit definitions of q_h^l, Q_h^l, and Q_l^l, we obtain

$$\frac{d\bar{\pi}^l}{d\nu} = \underbrace{\frac{\partial \bar{\pi}^l}{\partial q_h^l}}_{\leq 0} \underbrace{\frac{\partial q_h^l}{\partial \nu}}_{\leq 0} + \underbrace{\frac{\partial \bar{\pi}^l}{\partial Q_h^l}}_{\leq 0} \underbrace{\frac{\partial Q_h^l}{\partial \nu}}_{\leq 0} + \underbrace{\frac{\partial \bar{\pi}^l}{\partial Q_l^l}}_{\geq 0} \underbrace{\frac{\partial Q_l^l}{\partial \nu}}_{\geq 0} \geq 0,$$

where $\partial \bar{\pi}^l / \partial Q_h^l \leq 0$ follows from $\bar{\pi}^l > 0$. As ν approaches 1, the quantities q_h^l, Q_h^l, and Q_l^l become efficient, i.e., $\bar{\pi}^l = \bar{\pi}_l^*$. As ν approaches 0, φ approaches 0, and the quantities q_h^l and Q_h^l rise without bound, whereas Q_l^l becomes 0. This implies that starting from $\nu = 1$, where $\bar{\pi}^l = \bar{\pi}_l^*$, the value $\bar{\pi}^l$ decreases when lowering ν and there is some cutoff level $\bar{\nu}_l \in (0, 1)$ where $\bar{\pi}^l$ equals 0 and is negative below $\bar{\nu}_l$. Q.E.D.

References

Acemoglu, D., A. Makhdoumi, A. Malekian, and A. Ozdaglar (2019). "Too much data: prices and inefficiencies in data markets." Working Paper 26296. National Bureau of Economic Research.

Acquisti, A., and H. Varian (2005). "Conditioning prices on purchase history". Marketing Science 24(3), 367–381.

Ali, S., G. Lewis and S. Vasserman (2019). "Voluntary disclosure and personalized pricing." Technical report. National Bureau of Economic Research.

Battaglini, M. (2005). "Long-Term Contracting with Markovian Consumers", American Economic Review 95, 637–658

Bergemann, D., A. Bonatti, and T. Gan. (2019). "The economics of social data." Cowles foundation discussion paper 2203. Yale University.

Bergemann, D., B. Brooks, S. Morris (2015). "The Limits of Price Discrimination." American Economic Review 105 (3), 921–57.

Bonatti, A. and G. Cisternas (2020). "Consumer scores and price discrimination." Review of Economic Studies 87, 750–791.

Börgers, T., A. Hernando-Veciana, and D. Krähmer (2013). "When are Signals Complements or Substitutes?" Journal of Economic Theory 148, 165–195.

Brunnermeier, M., R. Lamba, and C. Segura-Rodriguez (2021). "Inverse selection", mimeo Princeton University

Choi, J.P., J. DohsShin, and B.C. Kim (2019). "Privacy and personal data collection with information externalities." Journal of Public Economics 173, 113–124.

Conitzer, V., C. Taylor, and L. Wagman (2012). "Hide and seek: costly consumer privacy in a market with repeat purchases." Marketing Science 31, 277–292.

Courty, P. and Li, H. (2000). "Sequential Screening.", The Review of Economic Studies 67, 697–717.

Crémer, J. and R. McLean (1985), "Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist when Demands are Interdependent." Econometrica, 53, 345–361.

Davis, A.S. (1961). "Markov Chains as Random Input Automata." The American Mathematical Monthly 68(3): 264–267.

De Cornière, A. and R. De Nijs (2016). "Online advertising and privacy." Rand Journal of Economics 47, 48–72.

De Cornière, A. and G. Taylor (2020). "Data and competition: a general framework with applications to mergers, market structure, and privacy policy." CEPR Discussion Paper No. DP14446.

Esö, P. and B. Szentes (2017). "Dynamic contracting: an irrelevance theorem." Theoretical Economics 12 (2017), 109–139.

Fudenberg, D. and J.M. Villas-Boas (2006). "Behavior-based price discrimination and customer recognition." Handbook on Economics and Information Systems 1, 377–436.

Hariri, Y. (2018). Homo Deus: A Brief History of Tomorrow. Harper Perennial.

Ichihashi, S. (2021). “Dynamic Privacy Choices.” mimeo Bank of Canada.

Ichihashi, S. and A. Smolin (2022), “Data Collection by an Informed Seller,” mimeo Toulouse School of Economics.

Laffont, J.J. and J. Tirole (1993). A Theory of Incentives in Procurement and Regulation. MIT Press.

Mussa, M and S. Rosen (1978). “Monopoly and Product Quantity.” Journal of Economic Theory 18, 301–317.

Riordan, M. and D. Sappington (1988). “Optimal Contracts with Public Ex Post Information.” Journal of Economic Theory 45, 189–199.

Strausz, R. (2009). “Entrepreneurial Financing, Advice, and Agency Costs” Journal of Economics & Management Strategy 18, 845–870.

Vives, X. and Z. Ye (2021). “Information Technology and Bank Competition,” mimeo IESE Business School.