Discussion Paper Series – CRC TR 224

Discussion Paper No. 442
Project B 03

Strategic Incentives and the Optimal Sale of Information

Rosina Rodríguez Olivera [1]

July 2023

[1] University of Bonn, Email: rrodrigu@uni-bonn.de

Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Collaborative Research Center Transregio 224 - www.crctr224.de
Rheinische Friedrich-Wilhelms-Universität Bonn - Universität Mannheim

Strategic Incentives and the Optimal Sale of Information

Rosina Rodríguez Olivera*

Abstract

I consider a model in which a monopolist data-seller offers information to privately informed data-buyers who play a game of incomplete information. I characterize the data-seller's optimal menu, which screens between two types of data-buyers. Data-buyers' preferences for information cannot generally be ordered across types. I show that the nature of data-buyers' preferences for information allows the data-seller to extract all surplus. In particular, the data-seller offers a perfectly informative experiment to the data-buyer with highest willingness to pay and a partially informative experiment, which makes the data-buyer with the highest willingness to pay for perfect information indifferent between both experiments. I also show that the features of the optimal menu are determined by the interaction between data-buyers' strategic incentives and the correlation of their private information. Namely, the data-seller offers two informative experiments even when data-buyers would choose the same action without supplemental information if data-buyers: i) have coordination incentives and their private information is negatively correlated or ii) have anti-coordination incentives and their private information is positively correlated.

JEL: D80, D82
Keywords: Screening, Information, Strategic incentives.

*University of Bonn, rrodrigu@uni-bonn.de. I thank David Miller, Anne-Katrin Roesler and Tilman Börgers for continued guidance and feedback. I also thank Laura Doval, Juan Dubra, Mathijs Jannsen, Daniel Krähmer, Louis-Pierre Lepage, Heng Liu, Franz Ostrizek, Dezsö Szalay, Vijay Subramanian and Ina Taneva for helpful discussions, as well as seminar participants at the University of Michigan and University of Bonn for their comments. Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project B03) is gratefully acknowledged.

The presence of firms who collect, aggregate, and sell information allows agents to supplement their private information and improve their decision making. Agents who buy information, data-buyers, then generally interact with each other in a market. For instance, firms acquire information about demand to guide their entry decisions and compete with others. Similarly, investors purchase information about the profitability of investments to choose whether or not to invest and may have incentives to coordinate. Since individual information decisions affect equilibrium outcomes, demand for information and therefore the optimal information offering of a data-seller depend on the strategic incentives between data-buyers.

In this paper, I analyze the direct sale of supplemental information in a stylized game of incomplete information. A data-seller owns a database containing information about a binary payoff-relevant state. The data-seller offers information to two privately informed data-buyers who play a two-stage game of incomplete information. In the information stage, data-buyers can simultaneously acquire supplemental information to reduce their uncertainty about the state. In the action stage, each data-buyer simultaneously selects an action from a binary set to maximize her expected payoff. The key feature is that the existence of private information makes the data-seller uncertain about demand for information and interacts with the strategic incentives of data-buyers, who must make inferences about the private information of others.

The model is motivated by applications such as the sale of information on debt monitoring by credit bureaus, which sell information on the credit worthiness of consumers and businesses to prospective lenders. An important consideration is that the information provided by such data-sellers is typically hard to gather and highly sensitive, so they have incentives to retain control of it by restricting access, for example to clients with a "permissable purpose" like banks. Moreover, in practice, it is often only feasible for data-sellers to condition pricing on their information offering, such as access to different versions of their database.[1] Accordingly, I assume that the data-seller is restricted to offering menus of information structures in which an item from the menu is self-selected by data-buyers, without requiring them to truthfully report their private information. Furthermore, I focus on a setting where data-buyers' information acquisition decisions are unobservable to other data-buyers, signal realizations from the data acquisition are private as well as conditionally

[1] Even if it is feasible to condition pricing on additional features such as another data-buyers' reports or their actions, such coordination would risk going against antitrust laws when data-buyers compete downstream.

independent, and only the information structure itself is contractible. These restrictions allow me to answer the following questions by isolating the impact of strategic incentives on optimal information provision in a restrictive but realistic class of mechanism. First, what is the data-seller's optimal menu of information offerings with multiple privately informed data-buyers? Second, how does it depend on the strategic incentives of data-buyers and the correlation between their private information?

Data-buyers' willingness to purchase supplemental information from the data-seller is determined by the precision and correlation of their private information, as well as strategic incentives in the action stage. The precision of private information determines overall demand for supplemental information, while its correlation impacts beliefs about the information observed by others, and therefore willingness to pay for information based on strategic incentives in the action stage. Data-buyers have *coordination incentives* (*anti-coordination incentives*) if the expected gain of choosing an action increases (decreases) in the probability that the other data-buyer chooses the same action. Accordingly, with coordination incentives (anti-coordination incentives), the willingness to pay of a data-buyer for information increases (decreases) in the precision of the information observed by others.

Data-buyers' private information induces two possible interim beliefs, interpreted as their type. When the state is also binary, types are one-dimensional and characterized by the probability that they assign to a given state. The "high type" is defined as the one that attaches a higher value to the fully informative experiment. I restrict attention to settings in which the value of information increases in its precision, so that the fully informative experiment is also the most valuable. The data-seller designs a personalized menu of Blackwell experiments and prices to screen data-buyer types, distorting the information provided to the low type in order to charge higher prices to the high type.

The optimal menu satisfies two standard properties of the screening literature: "no distortion at the top" and "no rent at the bottom". However, the data-seller can also extract all surplus from the high type. The full surplus extraction result arises from the nature of data-buyers' preferences for information. Since information is valuable to a data-buyer if and only if it affects their choice, data-buyer preferences for information depend on its precision (quality) and what the information is about (position). In fact, different types value information differently and may disagree on the ranking of partially informative experiments, implying that willingness to pay for information cannot be ordered across types. The data-seller captures all surplus by selecting the position of the information provided to the low type such that the high type is indifferent between both experiments.

Furthermore, the optimal menu is symmetric if the asymmetry of payoffs is sufficiently small. Otherwise, the data-seller offers no information to one data-buyer and a menu of information offerings to a single data-buyer. The main features of the optimal symmetric menu are as follows. The optimal menu contains the perfectly informative experiment offered to the high type and a concentrated experiment designed for the low type. The complete characterization of the information offered to the low type is determined by the distribution of data-buyer types and strategic incentives in the action stage. Consistent with previous work, if data-buyer types would take different actions without supplemental information (non-congruent beliefs), their preferences over partially informative experiments are not aligned (Bergemann et al., 2018). Accordingly, the data-seller offers partial information to the low type without attracting the high type. In contrast with previous work, if data-buyer types would choose the same action without supplemental information (congruent beliefs), I show that the data-seller still offers partial information to the low type when: (i) data-buyers have coordination incentives in the action stage and their private information is negatively correlated, or (ii) data-buyers have anti-coordination incentives and their private information is positively correlated. Further, for both congruent and non-congruent beliefs, I show that the quantitative properties of the optimal menu are determined by the interaction between strategic incentives and the correlation of private information.

Intuitively, the data-seller offers partial information to the low type when beliefs are congruent whenever the interaction between strategic incentives and the correlation of private information increases the demand for information for the low type. When private information is negatively correlated, data-buyers assign a higher probability to observing different private information. Hence, demand for supplemental information increases for the low type when data-buyers have coordination incentives, since it increases the correlation between their action choices. Similarly, when their private information is positively correlated and they have anti-coordination incentives, acquiring conditionally independent information is valuable, because it allows them to reduce the correlation between their actions.

These results highlight that the interaction between strategic incentives and the correlation of private information determines the features of the optimal menu. This interaction can relax the incentive compatibility constraints and expand the opportunity of the data-seller to serve both segments of the market, increasing profits. My results emphasize the importance of considering strategic interactions when designing information offerings, given that data-buyers generally interact with others. They also extend to a setting with N data-buyers in which payoffs depend on whether they match the state and the choice of the majority and to a setting in which the nature of coordination incentives are state-dependent. Lastly, I discuss the effect of allowing correlated signal realizations and a continuum of types on the optimal information provision.

Overall, my results highlight the importance of considering a setting in which multiple data-buyers interact, not only because of strategic incentives and the correlation of private information, but also because optimal information provision can depend on the distribution of buyer types in a market and the correlation between their signal realizations.

This paper contributes to the literature on information design in games with privately informed players and to the literature on selling information. In contrast to information design papers in which players have common priors (Taneva (2019), Mathevet et al. (2020)), I consider the role of private information in determining the seller's optimal information offering. Data-buyers can have heterogeneous previous experiences which provide private information about the state, affecting their demand for information and incentives for the data-seller to offer information. Within the information acquisition literature, previous work also studies information acquisition with multiple receivers who engage in strategic interactions, but have no initial private information (Admati and Pfleiderer (1986), Admati and Pfleiderer (1990), Hellwig and Veldkamp (2009), Myatt and Wallace (2011), Yang (2015), Amir and Lazzati (2016) and Kastl et al. (2018)). My results highlight that adding private information to these models has key implications for the optimal sale of information.

In relationship to the literature on information selling with private information, this paper is most closely related to Bergemann et al. (2018), Bonatti et al. (2022) and Bonatti et al. (2023). Bergemann et al. (2018) studies the design and ex-ante pricing of Blackwell experiments for a single privately informed receiver. In contrast, I consider a setting with multiple data-buyers, allowing me to study how the interplay between private information and strategic interactions affects the optimal information offering. I show that the structure of the optimal menu bears resemblance with the single agent case. However, the features of the optimal menu are determined by the interaction between the strategic incentives of data-buyers and the correlation of their private information. In particular, this interaction relaxes the incentive-compatibility constraints of data-buyers in comparison to the single agent case and allows the data-seller to offer at least partial information to both data-buyer types even when they have congruent beliefs. Furthermore, my analysis sheds light on the precise conditions under which some results from the single-buyer case extend to an analysis with multiple buyers. For example, previous results for a continuum of buyer types do not generalize to a setting with multiple buyers.Lastly, some of my analyses focus on other features that are inherently absent from a setting with a single buyer, such as correlation between signal realizations acquired by different buyers, but are relevant for a number of applications. Bonatti et al. (2022) also studies optimal mechanisms for a class of games with binary actions and states. In their setting, buyers' private types capture their marginal valuations and reveal nothing about the state of the world. In contrast, in my paper, buyers' types correspond to the realization of privately observed exogenous information about the state. Bonatti et al. (2023) analyzes symmetric games with quadratic payoffs, including both games of strategic substitutes and complements and fully characterize the optimal Gaussian mechanisms. As such, they restrict attention to linear-quadratic setting with Gaussian signals, but in a model with non-binary state and actions.

It is also related to Bergemann and Bonatti (2015), which studies a setting in which the data-seller engages in ex-post pricing and offers signal realizations instead. Eső and Szentes (2007) and Li and Shi (2017) also consider settings in which the data-seller engages in ex-post pricing, but in multi-player settings in which data-buyers' actions are contractible. Instead, I restrict attention to ex-ante pricing in which actions are not contractible. Kolotilin et al. (2017), Krähmer (2020), Candogan and Strack (2021), Segura-Rodriguez (2021), Yamashita and Zhu (2021) and Zhu (2021) consider a designer who selects a mechanism to provide supplemental information to multiple privately informed agent(s) based on their reported types. In contrast, I analyze a different contracting protocol in which prices are contingent only on the information itself and not the agents' actions nor the state. In this context, type-contingent information disclosure allows the data-seller to screen the data-buyers' types and increase its revenue. The assumptions I make about the contracting environment allow me to study the interaction of private information and strategic incentives in a similar setting to how information is sold in some key practical applications.[2] Lastly, Yang (2022) also characterizes the revenue-maximizing mechanism for data-seller who sells information, but this information is about consumer's preferences and sold to a privately informed producer. The producer is the unique data-buyer in this setting and it is privately informed of their costs, such that its private information is an inherent part of its preferences. In my setting, I characterize the revenue-maximizing menu of Blackwell experiments when data-buyers are privately informed about their prior belief about the state.

The remainder of the paper is organized as follows: Section 1 outlines the model, Section

[2]As discussed in Bergemann and Bonatti (2019), settings in which the price of information is only contingent on the information itself (and not its realization) captures more appropriately information products such as data appends, whereas settings in which the information is also contingent in actions and/or signal realizations represent information products such as marketing lists.

2 derives preliminary results, Section 3 characterizes the optimal menu, Section 4 discusses an application of my model and results, Section 5 studies extensions, including the case of N data-buyers, correlated signals, state-dependent coordination incentives and continuous types. Lastly, Section 7 concludes.

1 Model

Consider a setting with two data-buyers and one data-seller. Data-buyers play a game of incomplete information, where i indexes a generic data-buyer and j denotes the other. The payoff-relevant state ω is drawn from a binary set $\Omega = \{\omega_1, \omega_2\}$. Each data-buyer is privately informed about the state and attaches probability $\theta \in \{\theta_L, \theta_H\}$ to state ω_1. The correlation between data-buyers' private information is characterized by ρ and ν, where $\rho \in (0, 1)$ represents the probability that both data-buyers attach probability θ_H to state ω_1, whereas $\nu \in \left(0, \frac{1-\rho}{2}\right)$ represents the probability that data-buyers attach different probabilities to state ω_1.[3] The joint distribution of data-buyers' private information is displayed in Table 1.

Data-buyer i's type \ Data-buyer j's type	θ_L	θ_H
θ_L	$1 - 2\nu - \rho$	ν
θ_H	ν	ρ

Table 1: Joint distribution of private information.

The game has two stages: the information stage and the action stage. In the information stage, before the state ω is realized, the data-seller offers a personalized menu of Blackwell experiments about the state and prices to the data-buyers.[4] After the state ω is realized, each data-buyer observes her private information and data-buyers simultaneously decide whether or not to purchase a Blackwell experiment from the menu and if so, which one to acquire. If a data-buyer purchases information, she observes a private signal realization and updates her belief accordingly. Data-buyers don't observe each others' choices from

[3]This can be interpreted as data-buyers sharing a common prior and privately observing either good or bad news about the likelihood of state ω_1. Let $\mu_0 = \mathbb{P}(\omega = \omega_1)$ be their common prior and assume they observe a conditionally independent signal $s_0 \in \{s_0^1, s_0^2\}$ where $\mathbb{P}(s = s_0^1|\omega = \omega_k) = \mu_k$ with $k \in \{1, 2\}$. Then, $\rho = \mu_0(1-\mu_1)^2 + (1-\mu_0)(1-\mu_2)^2$ and $\nu = \mu_0(1-\mu_1)\mu_1 + (1-\mu_0)(1-\mu_2)\mu_2$.

[4]A Blackwell experiment provides information about the state, but not about the private information of the other data-buyer.

the menu.[5] In the action stage, each data-buyer simultaneously selects her action from the binary set $A = \{a_1, a_2\}$ to maximize her expected payoff conditional on her signal realization. The payoffs $u : A \times \Omega \to \mathbb{R}$, defined in Table 2, are symmetric and characterized by $c > 0$.[6]

$\omega = \omega_1$	a_1	a_2
a_1	$1, 1$	$c, 0$
a_2	$0, c$	$0, 0$

$\omega = \omega_2$	a_1	a_2
a_1	$0, 0$	$0, c$
a_2	$c, 0$	$1, 1$

Table 2: Action stage payoffs.

Under these assumptions, it is an ex-post dominant strategy for each data-buyer to match the state ω. The payoff parameter c determines a data-buyer's preference over the action of the other data-buyer. In particular, data-buyer i prefers when j selects the same action (a different action) when $c < 1$ ($c > 1$). Formally, data-buyers are said to have coordination (anti-coordination) incentives if the expected gain of choosing an action increases (decreases) in the probability that the other data-buyer chooses the same action. That is, data-buyers have coordination incentives if $c < 1$ and anti-coordination incentives if $c > 1$.[7] Note that the (anti-)coordination incentives are stronger the father away c is from $c = 1$.

Experiments. An individual experiment $E_i^m = (S^{m,i}, \{\pi^{m,i}(\cdot|\omega)\}_{\omega\in\Omega})$ provides data-buyer $i \in \{1, 2\}$ with information about the state ω and consists of a finite set of signal realizations $s_\ell^{m,i} \in S^{m,i}$ and a family of conditional distributions $\pi^{m,i}$ where

$$\pi_{\ell,k}^{m,i} := \mathbb{P}(s_\ell^{m,i}|\omega_k), \ \pi_{\ell,k}^{m,i} \geq 0 \text{ and } \sum_{\ell=1}^{L^{m,i}} \pi_{\ell,k}^{m,i} = 1$$

[5] Data-buyer i's deviations in information choices are unobservable, implying that action and information choices are strategically simultaneous.

[6] Any 2×2 symmetric game across players and state in which players strictly prefer to match the state can be normalized in this manner.

[7] Let I_i be data-buyer i's information set. Define σ_k as the probability that i assigns to j selecting a_1 conditional on state ω_k and I_i. Data-buyer i's expected gain of choosing action a_1 instead of a_2 conditional on her information set I_i, ΔU_i, is given by

$$\Delta U_i := \mathbb{P}(\omega = \omega_1|I_i)[\sigma_1 + (1-\sigma_1)c] - (1 - \mathbb{P}(\omega = \omega_1|I_i))\,[\sigma_2 c + (1-\sigma_2)]$$

where $\frac{\partial \Delta U_i}{\partial \sigma_k} \geq 0$ if and only if $c \leq 1$ for all k. That is, i's expected gain from selecting action a_1 instead of a_2 increases in the probability of j choosing action a_1 if and only if $c \leq 1$. Analogously, i's gain of choosing a_2 instead of a_1 increases in the probability that j chooses a_2 if and only if $c \leq 1$. See Taneva (2019) for more details.

with $L^{m,i} = |S^{m,i}|$. Denote by $\mathcal{E}_i$ the set of feasible experiments for data-buyer i. The seller's cost of providing information is zero.

An experiment E_i^m can be represented by a stochastic matrix in which each column represents a state and each row a signal realization, as in Table 3.

	ω_1	ω_2
$s_1^{m,i}$	$\pi_{1,1}^{m,i}$	$\pi_{1,2}^{m,i}$
$s_2^{m,i}$	$\pi_{2,1}^{m,i}$	$\pi_{2,2}^{m,i}$
$\vdots$	$\vdots$	$\vdots$
$s_{L^{m,i}}^{m,i}$	$\pi_{L^{m,i},1}^{m,i}$	$\pi_{L^{m,i},2}^{m,i}$

Table 3: Matrix representation of experiment E^m.

Assume that the realizations of data-buyers' private information and the realization of the signal $s_i \in S^{m,i}$ from any experiment E_i^m are independent conditional on the state ω. Moreover, assume that signal realizations between data-buyers are conditionally independent. The first assumption implies that the value of an experiment is determined by a data-buyer's private information, its correlation with the private information observed by others, and the nature of the strategic incentives. It also implies that the value of an experiment can be derived independently of its price. The second assumption rules out that signals can be used as a coordination device, except through their correlation with the state. As such, data-buyers attach no value to the uninformative experiment.

Data-seller's strategy space. The data-seller offers a menu of individual Blackwell experiments and prices with arbitrarily informative signals. Let $\mathcal{M} = (\mathcal{M}_1, \mathcal{M}_2)$ denote the menu of experiments offered by the data-seller. The menu $\mathcal{M}_i = (E_i^m, t_i^m)_{m\in\{1,\ldots,M\}}$ is offered to data-buyer $i \in \{1,2\}$, where experiment $E_i^m \in \mathcal{E}_i$ is offered to data-buyer i type θ_i at price $t_i^m \in \mathbb{R}$ and M is the number of experiments included in the menu with $m \in \{1, 2.., M\}$. Only the experiment itself is contractible, not its realization, the realized state, or the data-buyers' actions. Formally, a strategy for the data-seller is a menu $\mathcal{M} = (\mathcal{M}_1, \mathcal{M}_2)$, where $\mathcal{M}_i = (E_i^m, t_i^m)_{m=1}^M$, $t_i^m \in \mathbb{R}$ and $E_i^m \in \mathcal{E}_i$.

Data-buyer's strategy space. Each data-buyer i of type θ decides whether to supplement her private information. Let $\iota_{i\theta} \in \{0, 1, ..., M\}$ denote data-buyer i's information acquisition decision, where $\iota_{i\theta} = 0$ represents the case in which i doesn't acquire supplemental information and $\iota_{i\theta} = m$ the case in which i acquires experiment E_i^m. Conditional on all her information, data-buyer i chooses an action from the set $\{a_1, a_2\}$. Formally, a pure strategy for data-buyer i of type θ consists of a pair $(\iota_{i\theta}, \alpha_{i\theta})$ where $\iota_{i\theta} \in \{0, 1, ..., M\}$ and $\alpha_{i\theta} = (\alpha_{i,\iota_{i\theta}} : S^{\iota_{i\theta},i} \to \{a_1, a_2\})_{\iota_{i\theta}=0}^{M}$.

Solution concept. The solution concept is the data-seller's preferred perfect extended Bayesian equilibrium.[8] An equilibrium is an extended assessment satisfying consistency of beliefs, Bayesian updating, and sequential rationality in each information set. That is, conditional on a menu and information choices, each data-buyer i's action choice maximizes her expected payoff. Given a menu, data-buyer i's information choice maximizes the difference between her expected payoff in the action state and the price of information. Lastly, the optimal menu for the data-seller is the one that maximizes their expected profits, anticipating data-buyers' equilibrium choices given the equilibrium selection. In this context, the data-seller has the ability to design any statistical experiment for each profile of data-buyer types taking into account that information is an input into the data-buyers' strategic interaction which is beyond her control.

Definition 1 *A strategy profile (ι^*, α^*), a menu $\mathcal{M}^*$ and a belief system μ form an equilibrium if:*

i) (ι^*, α^*) and $\mathcal{M}^*$ satisfy sequential rationality. That is:

(a) Given $\mathcal{M}$,

$$\mathbb{E}[U_{i\theta}(\iota, \alpha^*)] \geq \mathbb{E}[U_{i\theta}(\iota, (\alpha_i', \alpha_{-i}^*))] \tag{1}$$

for all α_i' and $\iota \in \{0, ..., M\}^2$ and

$$\iota_{i\theta}^* \in \underset{\iota_{i\theta} \in \{0,...,M\}}{\arg\max} \; \mathbb{E}[U_{i\theta}((\iota_{i\theta}, \iota_{-i}^*), \alpha^*)] - t_i^{\iota_{i\theta}} \tag{2}$$

where expectations are taken over the state ω, the private information of the other data-buyer, and her choices.

(b) Let $\mathrm{E}(\mathcal{M})$ be the set of equilibria in the ensuing game after menu $\mathcal{M}$ is offered and denote by $e(\mathcal{M}) := (\iota^*, \alpha^*, \mu)$ a typical element of $\mathrm{E}(\mathcal{M})$. A menu $\mathcal{M}^*$ is optimal if is the solution to

$$\max_{\mathcal{M}} \max_{e(\mathcal{M}) \in \mathrm{E}(\mathcal{M})} \sum_{i \in \{1,2\}} \sum_{\theta \in \Theta} \sum_{m=1}^{M} \mathbb{P}(\theta_i = \theta) \mathbb{P}(\iota_{i\theta}^* = m | e(\mathcal{M})) \cdot t_i^m(e(\mathcal{M})).$$

[8] This definition is equivalent to weak Perfect Bayesian Equilibrium with the additional assumption that data-buyers do not update their beliefs about the state after observing a deviant menu. Since the data-seller chooses a menu before the state is realized, strategic independence only requires this additional constraint. See Battigalli (1996) or Watson (2016) for details.

ii) μ satisfies extended Bayesian updating.

iii) μ satisfies strategic independence: data-buyers don't infer anything about the state if the data-seller offers a deviant menu.

The value of information. The expected value of experiment E_i^m is defined as the marginal value of information, which corresponds to the difference in expected equilibrium payoffs with and without observing experiment E_i^m while fixing the equilibrium information acquisition of others. Denote by $V_{\mathcal{M}}(E_i^m;\theta)$ data-buyer i's expected value of experiment E_i^m when her interim belief is θ and the data-seller offers menu $\mathcal{M}$. Formally,

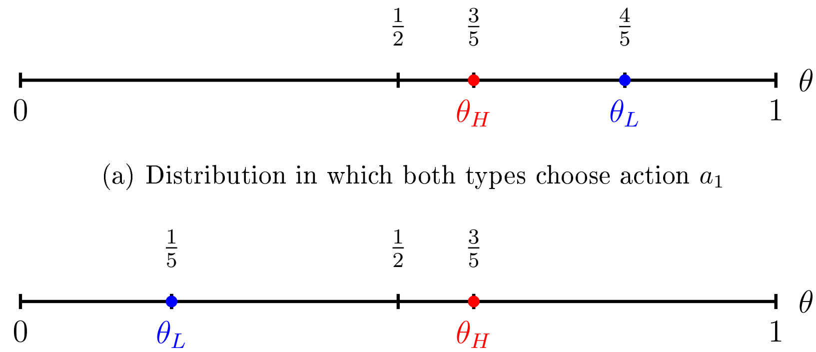
$$V_{\mathcal{M}}(E_i^m,\theta) = \mathbb{E}[U_{i\theta}((m,\iota_{-i}^*),\alpha^*)] - \mathbb{E}[U_{i\theta}((0,\iota_{-i}^*),\alpha^*)].$$

The value of experiment E_i^m for an individual data-buyer depends on her private information, her belief about the private information of others, strategic incentives in the action stage, and the menu. It is determined by the probability of matching the state, the probability of matching the action of the other data-buyer, and the payoff structure. The likelihood of matching the state depends on i's private information, but is independent of the information acquisition decision of the other data-buyer. In contrast, the likelihood of matching data-buyer j's action depends on j's private information. Lastly, the strategic environment in the action stage determines the preferences of data-buyers over equilibrium outcomes.

Let $\overline{E}$ denote the perfectly informative experiment. The high type θ_H is defined as the type that assigns higher value to $\overline{E}$. That is, $V_{\mathcal{M}}(\overline{E},\theta_H) \geq V_{\mathcal{M}}(\overline{E},\theta_L)$. For example, if $c=1$, data-buyers prefer action a_k if they assign a higher probability to state ω_k for $k\in\{1,2\}$ and they are indifferent between actions a_1 and a_2 if they assign equal probability to each state. In this case, the high type is the type closest to the cutoff $\frac{1}{2}$, as illustrated in Figure 1.

Lemma 1 provides a formal characterization of the high type θ_H. It states that the definition of the high type requires her to be sufficiently uncertain about the state in comparison to the low type.

Lemma 1 *Consider two types (θ_1,θ_2). Fix θ_1 and let $\hat{\theta}(c,\theta_1)$ be the data-buyer type that is initially indifferent between selecting action a_1 and a_2. If $\theta_2 \geq \hat{\theta}(c,\theta_1)$, there exists $\overline{\theta}(c,\theta_1)$ such that $V_{\mathcal{M}}(\overline{E},\theta_2) \geq V_{\mathcal{M}}(\overline{E},\theta_1)$ if and only if $\theta_2 \leq \overline{\theta}(c,\theta_1)$. Similarly, if $\theta_2 < \hat{\theta}(c,\theta_1)$, there exists $\underline{\theta}_H(c,\theta_1)$ such that $V_{\mathcal{M}}(\overline{E},\theta_2) \geq V_{\mathcal{M}}(\overline{E},\theta_1)$ if and only if $\theta_2 \geq \underline{\theta}(c,\theta_1)$. In both cases, we call θ_2 the high type and θ_1 the low type. Otherwise, θ_2 is called the low type and θ_1 the high type.*



(a) Distribution in which both types choose action a_1

(b) Distribution in which types choose different actions

Figure 1: Example of high (red) and low (blue) types for $c = 1$.

Figure 1 illustrates Lemma 1 for the case where $c = 1$. In both panels, $\hat{\theta}(c, \theta_1) = \frac{1}{2}$, $\theta_2 = \theta_H > \frac{1}{2}$ and $\theta_1 = \theta_L$. In panel a), $\overline{\theta}(1, \theta_1) = \theta_1$ whereas in panel b), $\overline{\theta}(1, \theta_1) = 1 - \theta_1$. Assume without loss of generality that the high type θ_H chooses action a_1 without supplemental information.

2 Preliminary results

2.1 Simplifications

The data-seller's problem can be simplified along two dimensions using two well-known results. First, the revelation principle of mechanism design implies that it is without loss of generality to focus on direct mechanisms in which the data-seller assigns one experiment to each data-buyer type θ. Second, the revelation principle of games of communication (Myerson, 1982) implies that it is without loss of generality to focus on experiments in which signals act as action recommendations . These two results imply that I can restrict attention to menus $\mathcal{M}_i$ with at most two elements ($M \leq 2$) for data-buyer i and experiments with two possible signal realizations ($S^{m,i} = \{s_1, s_2\}$ for all m and i).

Lemma 2 shows that the outcome of every menu can be attained by a direct menu which includes at most two elements for each data-buyer i. That is, the data-seller offers a menu which includes at most two experiments for each data-buyer i.

Lemma 2 *Let $\mathbb{M}^2 := \{(E_i^m, t_i^m)_{m \in \{1,\ldots,M\}} : \ M \leq 2 \text{ for all } i \in \{1, 2\}\}$ be the class of menus which include at most two experiments for each data-buyer i. Then, for any outcome that is attainable by any menu $\mathcal{M}$, there exists a menu $\mathcal{M}' \in \mathbb{M}^2$ that attains the same outcome.*

Given any direct menu $\mathcal{M}$, an experiment E_i^m with private signals is *responsive* if every signal $s \in S$ leads to a different action choice for data-buyer i of type θ. A direct menu $\mathcal{M}$ is responsive if every experiment $E_i^m \in \mathcal{M}_i$ is responsive for all i. Lemma 3 shows that it is without loss of generality to focus on menus in which the cardinality of the signal space equals the cardinality of the action space. I refer to these menus as responsive. Lemma 3 shows that the outcome of every direct menu can be attained by a responsive menu.

Lemma 3 *Let the class of responsive direct menus be*

$$\mathbb{M}^{2,2} := \{(E_i^m, t_i^m)_{m\in\{L,H\}} : \; S^{m,i} = \{s_1, s_2\} \text{ for all } m \in \{L, H\} \text{ and } i \in \{1, 2\}\}.$$

Then, for any outcome that is attainable by any menu $\mathcal{M} \in \mathbb{M}^2$, there exists an alternative menu $\mathcal{M}' \in \mathbb{M}^{2,2}$ that attains the same outcome.

Lemma 3 generalizes Proposition 1 from Bergemann et al. (2018) to a setting with multiple data-buyers. This result relies on two main assumptions: signals are private and information acquisition decisions are unobservable. This ensures that a change in the set of signals observed by one data-buyer has no effect on the other data-buyer's action choice. Hence, it is without loss of generality to consider $S^{m,i} = \{s_1, s_2\}$ and $\pi^{m,i} : \Omega \to [0, 1]^2$ for all $m \in \{1, ..., M\}$ and $i \in \{1, 2\}$. Then, an experiment E_i^m can be represented by the following matrix:

s/ω	ω_1	ω_2
s_1	$\pi_1^{m,i}$	$1 - \pi_2^{m,i}$
s_2	$1 - \pi_1^{m,i}$	$\pi_2^{m,i}$

where $(\pi_1^{m,i}, \pi_2^{m,i}) \in [0, 1]^2$ for all $m \in \{1, ..., M\}$ and i. Given that signals act as action recommendations, after observing signal s_k, data-buyer i must be willing to choose action a_k when j follows her action recommendation if she chooses to acquire supplemental information. That is, if i acquires experiment n and j acquires experiment m, then

$$\theta_i \pi_1^{n,i}[\pi_1^{m,j} + (1 - \pi_1^{m,j})c] \geq (1 - \theta_i)(1 - \pi_2^{n,i})[(1 - \pi_2^{m,j})c + \pi_2^{m,j}] \text{ and}$$
$$(1 - \theta_i)\pi_2^{n,i}[(1 - \pi_2^{m,j})c + \pi_2^{m,j}] \geq \theta_i(1 - \pi_1^{n,i})[\pi_1^{m,j} + (1 - \pi_1^{m,j})c]$$

Also, without loss of generality let the likelihood of observing signal s_1 is higher conditional on state ω_1 than ω_2 compared to s_2 for all i.

Assumption 1

$$\frac{\mathbb{P}(s = s_1|\omega_1)}{\mathbb{P}(s = s_1|\omega_2)} \geq \frac{\mathbb{P}(s = s_2|\omega_1)}{\mathbb{P}(s = s_2|\omega_2)} \Leftrightarrow \frac{\pi_1^{m,i}}{1 - \pi_2^{m,i}} \geq \frac{1 - \pi_1^{m,i}}{\pi_2^{m,i}} \Leftrightarrow \pi_1^{m,i} + \pi_2^{m,i} \geq 1.$$

These results generalize the ones from Bergemann et al. (2018) from a decision problem to a game setting and are intuitively related to Myerson (1982) and Taneva (2019). In particular, Taneva (2019) studies how to derive optimal information structures in static finite environments in which agents share a common prior and shows that it is without loss of generality to restrict attention to direct information structures, in which signals act as action recommendations. Lemma 3 extends this intuition to a setting in which agents have private information. As a result, the designer can design a menu of information offerings which screens for this private information.

2.2 The value of experiments

In this section, I derive a closed form expression for the value of experiments.[9] Assume that data-buyer j type θ_j follows her equilibrium strategy, and denote by m her experiment choice. Since signals act as action recommendations, data-buyer j type θ_j conditions her action choice on the realized signal and selects action a_k after observing signal s_k.[10] The value of experiment E_i^n compares data-buyer i's expected equilibrium payoff across two cases:

i) she acquires experiment E_i^n and selects action a_k after observing signal s_k,

ii) she acquires no supplemental information and selects either action a_1 or a_2.

In both cases, j follows her equilibrium strategy, acquiring experiment E_j^m and selecting action a_k after observing signal s_k.

Define $v_k(E_i^n, \theta_i; m)$ as data-buyer i's expected gain of acquiring experiment E_i^n if, without information, she would choose action a_k while j plays her equilibrium strategy. Data-buyer i's expected gain of acquiring information when choosing action a_2 and a_1 without supplemental information followed by action a_k after observing signal s_k, respectively, are given by:

$$v_2(E_i^n, \theta_i; m) = \underbrace{\theta_i \pi_1^{n,i} \left[\pi_1^{m,j} + \left(1 - \pi_1^{m,j}\right) c\right]}_{\mathbb{P}(\omega=\omega_1)\mathbb{P}(s^i=s_1|\omega=\omega_1)[\mathbb{P}(s^j=s_1|\omega=\omega_1)+\mathbb{P}(s^j=s_2|\omega=\omega_1)c]} - \overbrace{(1-\theta_i)\left(1 - \pi_2^{n,i}\right)\left[\left(1 - \pi_2^{m,j}\right) c + \pi_2^{m,j}\right]}^{\mathbb{P}(\omega=\omega_2)\mathbb{P}(s^i=s_1|\omega=\omega_2)[\mathbb{P}(s^j=s_1|\omega=\omega_2)c+\mathbb{P}(s^j=s_2|\omega=\omega_2)]} \text{ and}$$

$$v_1(E_i^n, \theta_i; m) = (1-\theta_i)\pi_2^{n,i}\left[\left(1 - \pi_2^{m,j}\right) c + \pi_2^{m,j}\right] - \theta_i\left(1 - \pi_1^{n,i}\right)\left[\pi_1^{m,j} + \left(1 - \pi_1^{m,j}\right) c\right].$$

That is, $v_k(E_i^n, \theta_i; m)$ is the difference between the data-buyer's expected gain in state ω_k' and her expected loss in state ω_k after observing signal s_k'. Then, data-buyer i's value of

[9] A data-buyer values an experiment if and only if she conditions her choice on the signal realization, such that signals act as action recommendations.

[10] Consistency of beliefs implies that $\mathbb{P}(\iota_j = m|\theta_j) = 1$ and $\mathbb{P}(a_j = a_k|\theta_j, \iota_j = m, s^j = s_k) = 1 \ \forall\ k \in \{1, 2\}$.

experiment E_i^n when the data-seller offers the menu $\mathcal{M}$ is given by:

$$V_{\mathcal{M}}(E_i^n, \theta_n) = \max\left\{0, \sum_{m\in\{L,H\}} \mathbb{P}(\theta_j = \theta_m|\theta_n) v_k(E_i^n, \theta_n; m) \text{ s.t. } k \text{ solves } \alpha_{i,\iota_{i\theta}=0} = a_k\right\}.$$

Combined with the definition of data-buyer types, I characterize the support of feasible distributions of data-buyer types according to Lemma 4.

Lemma 4 *Assume without loss of generality that the high type θ_H chooses action a_1 without supplemental information. Then, for all $c > 0$, the distribution of buyer-types satisfies*

$$\theta_H \geq \min\left\{\frac{1}{1+c}, \frac{c}{1+c}\right\},$$

i) $\theta_L \geq \min\left\{\frac{1}{1+c}, \frac{c}{1+c}\right\}$ and $\theta_H < \theta_L$ when θ_L also chooses a_1 without supplemental information.

ii) $\theta_L \leq \max\left\{\frac{1}{1+c}, \frac{c}{1+c}\right\}$ and $\theta_L < \theta_H < 1 - \theta_L$, otherwise.

2.3 The data-seller's problem

Denote by E_i^L and E_i^H the experiment that the data-seller designs for data-buyer i type θ_L and for θ_H, respectively. The presence of private information implies that the data-seller is uncertain about demand for experiments and must screen data-buyer types.[11] The data-seller's problem is then to design a menu of experiments to maximize expected transfers subject to data-buyers' incentive-compatibility and participation constraints.[12] That is:

$$\max_{(E_i^m, t_i^m)_{(m,i)\in\{L,H\}\times\{1,2\}}} (1-\nu-\rho)(t_1^L + t_2^L) + (\rho+\nu)(t_1^H + t_2^H)$$

subject to the participation constraints

$$IR_{L,i} : V_{\mathcal{M}}(E_i^L, \theta_L) - t_i^L \geq 0,\ IR_{H,i} : V_{\mathcal{M}}(E_i^H, \theta_H) - t_i^H \geq 0$$

and the incentive-compatibility constraints

$$IC_{L,i} : V_{\mathcal{M}}(E_i^L, \theta_L) - t_i^L \geq V_{\mathcal{M}}(E_i^H, \theta_L) - t_i^H$$
$$IC_{H,i} : V_{\mathcal{M}}(E_i^H, \theta_H) - t_i^H \geq V_{\mathcal{M}}(E_i^L, \theta_H) - t_i^L$$

[11]If data-buyers have no private information, the data-seller faces no uncertainty about the demand for experiments and finds it optimal to offer perfect information priced at the willingness to pay of the data-buyer.

[12]Payments are conditional only on the information product itself and not on the types of other data-buyers. This assumption rules out the use of the Cremer-McLean condition (Crémer and McLean, 1988).

for all $i \in \{1, 2\}$, where data-buyer i's value of experiment E_i^n when the data-seller offers the menu $\mathcal{M}$.

I restrict the data-seller to a menu mechanism. Every data-buyer simple selects an experiment, the informativeness and the transfer of her experiments doesn't depend on the choices of the other data-buyer. This restriction is inline with the motivating applications. Buyers of credit worthiness, for example, can choose the level of detail of the report they purchase, but the information and prices are independent of the purchasing behavior of other firms. From a technical point of view, this restriction rules out the use of Crémer and McLean (1988) and Krähmer (2020) mechanisms for full surplus extraction.

Benchmark: No private information. Consider as a benchmark the case in which data-buyers have no private information about the state, i.e., $\theta_L = \theta_H = \theta$. In this case, the data-seller's problem is to maximize her expected profits by selecting an experiment E_i for data-buyer i subject to the participation constraints. Specifically, it is given by

$$\max_{(E_i, t_i)_{i \in \{1,2\}}} (t_1 + t_2) \text{ s.t. } V(E_i, \theta) - t_i \geq 0$$

where

$$V(E_i, \theta) = \begin{cases} \max\{0, (1-\theta)\pi_2^i(c(1-\pi_2^j) + \pi_2^j) - \theta(1-\pi_1^i)(c(1-\pi_1^j) + \pi_1^j)\} & if\ \alpha_{i,\iota_{i\theta}=0} = a_1 \\ \max\{0, \theta\pi_1^i(c(1-\pi_1^j) + \pi_1^j) - (1-\theta)(1-\pi_2^i)(c(1-\pi_2^j) + \pi_2^j)\} & if\ \alpha_{i,\iota_{i\theta}=0} = a_2 \end{cases}$$

In both cases, it is optimal for the data-seller to set $t_i = V(E_i, \theta)$ and that the optimal experiments E_1 and E_2 depend on the value of c. In particular, when $c < 2$, it is optimal for the data-seller to offer the same information to both data-buyers and to offer perfect information. In contrast, when $c > 2$, it is optimal for the data-seller to offer personalized information and offer perfect information to one data-buyer and no information to the other.

Benchmark menus. Consider a set of four benchmark feasible menus which trivially satisfy the participation and incentive-compatibility constraints of the data-seller's problem while only including either full or no information. Lemma 5 shows that all of these menus guarantee the data-seller a strictly positive profit, providing a lower-bound on the profits achievable by an optimal menu.

Formally, let $\overline{E}$ and $\underline{E}$ denote the fully informative and uninformative experiment, respectively. There are four menus that satisfy all participation and incentive-compatibility constraints and allow the data-seller to obtain strictly positive profits. These menus are:

1. $\underline{\mathcal{M}}^1$ given by $E_i^k = \overline{E}$ and $t_i^k = V_{\underline{\mathcal{M}}^1}(\overline{E}, \theta_L)$ for all $i \in \{1,2\}$ and $k \in \{L, H\}$. That is, the data-seller can offer only the perfectly informative experiment to both data-buyers at a fixed price equal to the low type's willingness to pay.
2. $\underline{\mathcal{M}}^2$ given by $E_i^H = \overline{E}$, $E_i^L = \underline{E}$, $t_i^L = 0$ and $t_i^H = V_{\underline{\mathcal{M}}^2}(\overline{E}, \theta_H)$ for all $i \in \{1,2\}$. That is, the data-seller can offer the same menu to both data-buyers, which includes the fully informative experiment for the high type and no information to the low type.
3. $\underline{\mathcal{M}}^3$ given by $E_i^k = \overline{E}$, $E_j^k = \underline{E}$, $t_i^k = V_{\underline{\mathcal{M}}^3}(\overline{E}, \theta_L)$ and $t_j^k = 0$ for all $i \in \{1,2\}$, $j \neq i$, $k \in \{L, H\}$. That is, the data-seller can offer no information to both types of data-buyer j and the fully informative experiment to both types of data-buyer i priced at the low type's willingness to pay.
4. $\underline{\mathcal{M}}^4$ given by $E_i^H = \overline{E}$, $E_i^L = E_j^k = \underline{E}$, $t_i^L = 0$, $t_i^H = V_{\underline{\mathcal{M}}^4}(\overline{E}, \theta_H)$ and $t_j^k = 0$ for all $i \in \{1,2\}$, $j \neq i$, $k \in \{L, H\}$. That is, the data-seller can offer no information to both types of data-buyer j and to the low type of data-buyer i and the fully informative experiment to the high type of data-buyer i.

Lemma 5 *The data-seller can guarantee herself strictly profits by selecting any menu*

$$\underline{\mathcal{M}} \in \{\underline{\mathcal{M}}^1, \underline{\mathcal{M}}^2, \underline{\mathcal{M}}^3, \underline{\mathcal{M}}^4\}.$$

Optimality of a menu with two distinct experiments. Lemma 5 implies that the data-seller can guarantee herself a certain level of profits without screening data-buyer types, because she can always offer a menu which offers either 1) only the perfectly informative experiment offered to both data-buyers at a fixed price, p, equal to the low type's willingness to pay or 2) only the perfectly informative experiment to one data-buyer at a price equal to the low type's willingness to pay and no information to the other data-buyer. Accordingly, Lemma 6 shows that it is optimal for the data-seller to offer a menu with two distinct items to at least one data-buyer if and only if the probability of the data-buyer being the high type is sufficiently high.

Lemma 6 *It is optimal for the data-seller to offer a menu $\mathcal{M}$ which offers two distinct elements to at least one data-buyer if and only if $\nu + \rho \geq \frac{\widehat{\Pi} - \left(t_1^L + t_2^L\right)}{t_1^H + t_2^H - \left(t_1^L + t_2^L\right)}$, where*

$$\widehat{\Pi} = \begin{cases} (1-\theta_L)\max\{2, c\} & \text{if } a_1 \text{ is chosen without supplemental information by } \theta_L \\ \theta_L \max\left\{2, \left[\frac{(1-2\nu-\rho)}{1-\nu-\rho}c + \frac{\nu}{1-\nu-\rho}\right]\right\} & \text{if } a_2 \text{ is chosen without supplemental information by } \theta_L. \end{cases}$$

Note that when $c < 2$, the data-seller's guaranteed payoff is higher when she offers the same information to both data-buyers. Otherwise, the data-seller can guarantee herself a higher payoff by offering personalized information.

3 Optimal menu of experiments

3.1 General properties

The optimal menu shares some structural properties with the one data-buyer setting established in Bergemann et al. (2018). Proposition 1 generalizes these results to a two data-buyer setting and identifies which constraints are binding as well as the information provided to the high type in any optimal menu.

Proposition 1 *In an optimal menu:*

i) E_i^L *and* E_i^H *are concentrated, i.e,* $\pi_k^{m,i} = \mathbb{P}_{E_i^m}(s_i = s_k | \omega = \omega_k) = 1$ *for all* $m \in \{L, H\}$, $i \in \{1, 2\}$ *and for some* $k \in \{1, 2\}$.

ii) E_i^H *is fully informative for at least one data-buyer.*

iii) Both participation constraints bind for all i.

iv) The incentive-compatibility constraint of the high type binds for all i.

The optimal menu satisfies two standard properties of the screening literature: "no distortion at the top" and "no rent at the bottom". However, the data-seller can also extract all surplus from the high type, because information is valuable only when it affects data-buyers' decision making. As such, information has two relevant dimensions: its precision and its position.[13] All data-buyer types prefer experiments with higher precision. However, different data-buyer types may disagree on their preference for the position of information, given that they may select different actions if they don't supplement their private information. Therefore, preferences for information cannot be ordered across types. As a result, the data-seller captures all the surplus by selecting the position of information such that the high type is indifferent between the experiments offered.

Moreover, both types are offered an experiment in which the distribution of signals conditional on one state ω_k is degenerated, eliminating uncertainty for one state. Given that data-buyers have incentives to match the state, the data-seller can shift the probability mass from $1-\pi_k^{m,i}$ to $\pi_k^{m,i}$ with $k \in \{1, 2\}$ until one of them reaches 1. In particular, if the low type would choose action a_1 (a_2) without supplemental information, it is optimal to set $\pi_1^{L,i} = 1$

[13] For example, define $\pi_1^{m,i} + \pi_2^{m,i}$ as the precision of experiment E_i^m and $\pi_1^{m,i} - \pi_2^{m,i}$ as its position.

($\pi_2^{L,i} = 1$). Similarly, since the high type would choose action a_1 without supplemental information, it is optimal to set $\pi_1^{H,i} = 1$. Intuitively, the data-seller offers each data-buyer i of type θ_i an experiment that reveals without noise the state that matches the action that they would have selected without supplemental information.

Lastly, the perfectly informative experiment is part of any optimal menu and is offered to at least one data-buyer. It is the most valued by any buyer type, since it allows them to perfectly match the state. Hence, if this experiment is not part of a menu, the data-seller can replace the currently most informative experiment with the perfectly informative one, weakly increasing profits by charging a higher price for this experiment while ensuring that the incentive-compatibility constraints are satisfied. Furthermore, it is optimal for the data-seller to offer this experiment to at least one of the high types, since their willingness to pay is higher.

3.1.1 Optimality of symmetric menus

Proposition 2 shows that it is optimal for the data-seller to offer the same menu of Blackwell experiments to both data-buyers when c is sufficiently small. Intuitively, it is trivial that when data-buyers have coordination incentives ($c < 1$), data-buyers are better off when they are offered the same menu since this increases the likelihood of choosing the same action. When data-buyers have anti-coordination incentives, a reasonable conjecture is that the data-seller may find it profitable to offer information to only one data-buyer, since this increases the willingness to pay for information of this data-buyer. However, when c is sufficiently small, it is not profitable for the data-seller to offer a personalized menu. This is because the personalized menu would increase the willingness to pay of one data-buyer, but the data-seller would only serve one of them. In contrast, when c is large, the data-seller is better off serving only one of the data-buyers, offering a personalized menu.

Proposition 2 *The optimal menu is symmetric if:*

i) $c \leq \frac{2\rho}{\rho-\nu}$ *when data-buyers would choose the same action without supplemental information.*

ii) $c \leq \frac{2\rho+\nu}{\rho}$ *when data-buyers would choose different actions without supplemental information.*

Otherwise, the data-seller offers either:

i) a symmetric menu or,

ii) no information to one data-buyer and a menu containing full and no information for the other one or, *iii) no information to one data-buyer and a menu containing full and partial information for the other one or,*

iv) full information to one data-buyer and no information to the other.

3.2 Complete characterization

When the optimal menu is symmetric, the complete characterization of the information provided to the low type depends on the:

1. *Payoff environment*: The payoff environment determines the presence of coordination ($c < 1$) or anti-coordination incentives ($c > 1$), which pins down the effect of information observed by others on their willingness to pay for an experiment.
2. *Correlation between data-buyers' private information:* The correlation between data-buyers' private information affects their beliefs about the information observed by others. In particular, data-buyers' private information is positively (negatively) correlated if $\nu < \sqrt{\rho} - \rho$ ($\nu > \sqrt{\rho} - \rho$).
3. *Congruency of beliefs*: The support of the distribution of data-buyer types determines whether or not they choose the same action if they don't supplement their information, which impacts a data-buyer's ranking of partially informative experiments. Data-buyers' interim beliefs are congruent (non-congruent) if both types choose the same action (different actions) without supplemental information.

Lemma 4 characterizes the distributions of data-buyer types that satisfy congruency (non-congruency) of data-buyers' interim beliefs. This is formalized in Definition 2.

Definition 2 *Assume that the high type θ_H chooses action a_1 without supplemental information. Beliefs are congruent if*

$$i)\ \theta_L > \theta_H \geq \frac{c}{1+c} \text{ when } c < 1 \text{ and } ii)\ \theta_L > \theta_H \geq \frac{1}{1+c} \text{ when } c > 1$$

and non-congruent if

$$i)\ \theta_H \geq \frac{c}{1+c},\ \theta_L \leq \frac{1}{1+c} \text{ and } \theta_L < \theta_H < 1 - \theta_L \text{ when } c < 1$$
$$ii)\ \theta_H \geq \frac{1}{1+c},\ \theta_L \leq \frac{c}{1+c} \text{ and } \theta_L < \theta_H < 1 - \theta_L \text{ when } c > 1.$$

3.2.1 Optimal menu with non-congruent beliefs

When beliefs are non-congruent, the optimal experiment offered to the low type is partially informative, as stated in Proposition 3. Proposition 3 implies that the qualitative results from Bergemann et al. (2018) for the one data-buyer setting extend to multiple data-buyers. That is, the high type is offered the perfectly informative experiment and the low type is offered a partially informative experiment.

Proposition 3 *Suppose that beliefs are non-congruent. In an optimal menu, the data-seller offers perfect information to the high type and partial information to the low type, where the former is denoted by E^H and the latter by E^L.*

With non-congruent beliefs, different data-buyer types may disagree on the ranking of partially informative experiments, as illustrated in Figure 2. For instance, if $(\theta_L, \theta_H) = (0.2, 0.6)$, the low type prefers experiment $(1/2, 1)$ to experiment $(1, 1/2)$, whereas the high type prefers experiment $(1, 1/2)$ to $(1/2, 1)$. As a result, the data-seller can offer the low type information that has less value to the high type by making it sufficiently imprecise (π_1^L sufficiently low). In an optimal menu, the data-seller selects π_1^L such that the high type is indifferent between acquiring experiment E^L or E^H. That is,

$$t^H - t^L = \theta_H(1 - \pi_1^L)\left[\frac{\nu}{\nu + \rho}(\pi_1^L + (1 - \pi_1^L)c) + \frac{\rho}{\nu + \rho}\right].$$

The right-hand side is the product of the probability of state ω_1, the additional precision from acquiring information, and the gain of choosing action a_1 over action a_2 when the state is ω_1. The left-hand side is the price differential. Hence, the information offered to the low type is such that the price differential equals the expected gain in state ω_1.

Even though the qualitative properties of the optimal menu are independent of the strategic incentives and the correlation of private information, its quantitative properties are determined by their interaction, as stated in Lemma 7.

Lemma 7 *The precision of the optimal E^L decreases as coordination incentives increase. Moreover, the effect of increasing the correlation of private information depends on coordination incentives:*

i) If data-buyers have coordination incentives ($c < 1$), the precision of the optimal E^L decreases in the correlation of private information.

ii) If data-buyers have anti-coordination incentives ($c > 1$), the precision of the optimal E^L increases in the correlation of private information.

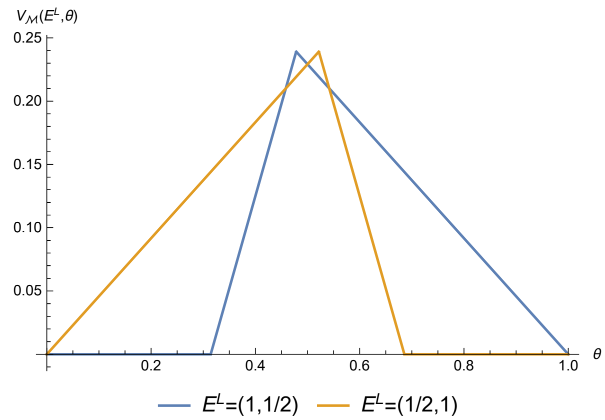


Figure 2: Value of E^L when private information is conditionally independent ($(\nu, \rho) = \left(\frac{1}{4}, \frac{1}{4}\right)$), data-buyers have coordination incentives ($c = \frac{2}{3}$) and E^H is fully informative.

The precision of the optimal experiment E^L decreases as coordination incentives increase, because the value of E^L increases for the high type and decreases for the low type. Then, an increase in coordination incentives implies that the high type has higher incentives to deviate, reducing the data-seller's scope to provide information to the low type. Moreover, with coordination (anti-coordination) incentives, the precision of E^L decreases (increases) in the correlation of private information.

3.2.2 Optimal menu with congruent beliefs

When data-buyers' beliefs are congruent, the information offered to the low type is determined by the interaction between strategic incentives in the action stage and the correlation of private information, as stated in Proposition 4.

Proposition 4 *Assume data-buyers' beliefs are congruent. In an optimal menu, there exists $\hat{\theta} \in (0, 1)$ such that the data-seller offers perfect information to the high type and offers the low type:*

i) partial information, if data-buyers' private information is negatively correlated, they have coordination incentives and $\theta_L < \hat{\theta}$.

ii) partial information, if data-buyers' private information is positively correlated, they have anti-coordination incentives and $\theta_L < \hat{\theta}$.

iii) no information, otherwise.

Strategic incentives play no role in determining the features of an optimal menu if and only if private information is conditionally independent ($\nu = \sqrt{\rho} - \rho$) or data-buyers' payoffs are independent of each others' choices ($c = 1$). In both cases, the qualitative properties of a one data-buyer menu generalize to a two data-buyer setting. That is, in any optimal menu, the high type learns the state and the low type is offered no information. Otherwise, the optimal menu is determined by the interaction between strategic interactions and the correlation of private information.

When there are coordination incentives ($c < 1$), data-buyers face no trade-off between matching the state and each others' actions. Hence, the value of an experiment increases in the precision of the experiment observed by others, because it increases the correlation between the state and their action choices, allowing data-buyer i to better predict j's action choice. Information acts as a coordination device and is valuable for two reasons: it reduces uncertainty about the state and about the choices of other data-buyers. When predicting the action choice of the other data-buyer, each data-buyer makes inferences about what information has been gathered by the other, which depends on the correlation between their private information. If their private information is positively (negatively) correlated, data-buyers assign a higher (lower) probability to observing the same private information and acquiring the same experiment. Thus, demand for information increases (decreases) for the low (high) type when private information is negatively correlated, because they assign a higher (lower) probability to encountering a high type which observes full information. This, in turn, implies that it is optimal for the seller to set a lower price for full information and increase the price for partial information. As a result, partial information is relatively less attractive for the high type, creating scope for the data-seller to offer some partial information to the low type, as long as the low type is sufficiently unsure about the state. When private information is positively correlated, demand for information by the low type is reduced in comparison to the conditionally independent case in which the low type is offered no information. As such, the low type is also offered no information.

When data-buyers have anti-coordination incentives ($c > 1$), the trade-off between matching the state and each others' actions implies that the value of experiment E^n decreases in the precision of the information observed by others, since it increases the correlation between their action and the state. Data-buyers want to be as informed as possible about the state, but their choices to be as uncorrelated with each other as possible. Information is still valuable because it allows data-buyers to learn about the state, but the low (high) type values information more (less) when their private information is positively correlated, because it reduces (increases) the correlation between their action choices. If the low type is sufficiently uncertain about the state, the increase in demand allows the data-seller to offer some partial information to the low type even when beliefs are congruent without incurring any cost in terms of surplus extraction from the high type. In contrast, acquiring supplemental information when private information is negatively correlated increases the correlation between action choices through increasing the correlation with the state, decreasing the low type's willingness to pay for supplemental information with respect to the conditionally independent case. As such, the data-seller also offers no information to the low type.

In both cases, the partially informative experiment offered to the low type is such that the price differential equals the expected gain in state ω_2. That is:

$$t^H - t^L = (1-\theta_H)(1-\pi_2^L)\left[1+\pi_2^L\left(\frac{\nu}{\nu+\rho}\right)(1-c)\right].$$

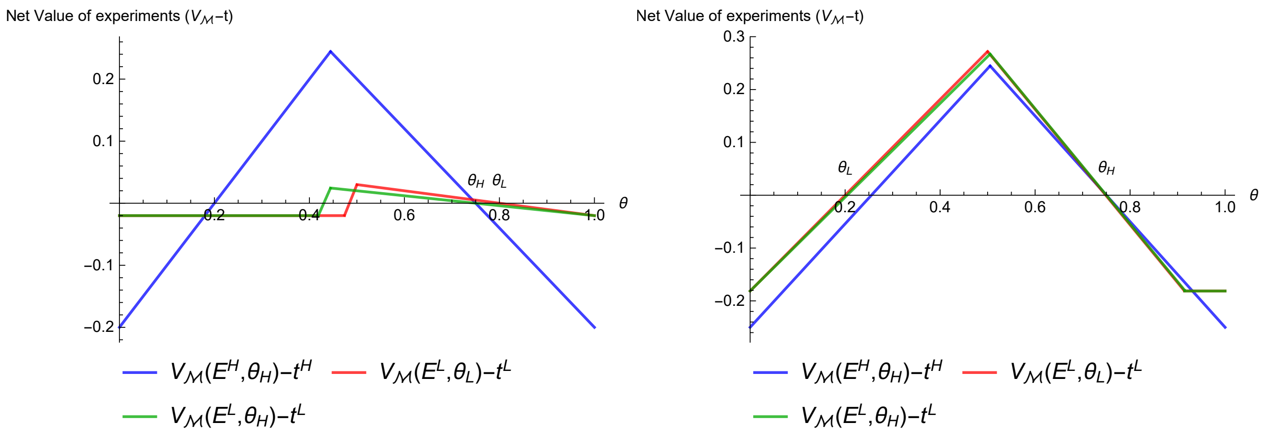
The left-hand side is the price differential. The right-hand side is the product of the probability of state ω_2, the probability of observing signal s_1 in state ω_2 and the expected payoff gain. The expected payoff gain depends on the probability of observing different private information, the coordination incentives and the precision of experiment E^1 in state ω_2.

Figure 3 shows an example of the optimal menus for two data-buyer type distributions, one with congruent beliefs and one with non-congruent beliefs, assuming the same strategic environment and correlation of private information.[14] It illustrates how the interaction between strategic incentives and correlated private information creates scope for the data-seller to offer partial information to the low type, even when data-buyers have congruent types. Both panels depict the value of two experiments, E^L and E^H, net of their prices, as a function of data-buyer type θ. In both panels, the optimal menu is full surplus extraction by the data-seller. This is shown by the intersection of $V_{\mathcal{M}}(E^L,\theta_L)-t^L$ and $V_{\mathcal{M}}(E^H,\theta_H)-t^H$ with the x-axis (net value of zero) at their type. Second, the net value of E^L for the high type ($V_{\mathcal{M}}(E^L,\theta_H)-t^H$) is also zero at θ_L, implying that the high type is indifferent between acquiring experiments E^L and E^H. Compared to previous work in the literature, $V_{\mathcal{M}}(E^L,\theta_H)$ and $V_{\mathcal{M}}(E^L,\theta_L)$ differ due to the correlation between data-buyers' private information.

Lastly, the quantitative properties of the optimal menu also depend on strategic incentives and the correlation of private information, as stated in Lemma 8. This result and its intuition are analogous to Lemma 7.

Lemma 8 *Assume that the optimal E^L is partially informative. The precision of the optimal*

[14] When data-buyers have congruent (non-congruent) beliefs, the optimal menu offers the perfectly informative experiment to the high type at a price of $t^H = 0.2$ ($t^H = 0.25$) and partial information to the low type, characterized by $\pi_1^L = 1$ and $\pi_2^L = 0.1$ ($\pi_1^L = 0.9$ and $\pi_2^L = 1$), at a price of $t^L = 0.05$ ($t^L = 0.18$).



(a) Congruent beliefs: $\theta_H = \frac{3}{4}$ and $\theta_L = \frac{4}{5}$

(b) Non-congruent beliefs: $\theta_H = \frac{3}{4}$ and $\theta_L = \frac{1}{5}$

Figure 3: Optimal menu with coordination incentives ($c = \frac{2}{3}$) and negatively correlated private information ($\rho = 0.2$ and $\nu = 0.4$).

E^L decreases as coordination incentives increase and decreases (increases) in the correlation of private information when data-buyers have coordination (anti-coordination) incentives.

3.2.3 Effect of congruency of beliefs on screening

The congruency of beliefs has a significant effect on the incentive-compatibility constraints of the data-buyers and on the ability of the data-seller to screen between their types. In particular, when data-buyers have congruent beliefs their private information yields similar interim beliefs. In contrast, when data-buyers have non-congruent beliefs, their private information is sufficiently different to induce a sufficiently high difference in interim beliefs such that different actions are optimal without supplementing their information. This difference implies that screening the types is easier in the non-congruent case.

This can be understood in the spirit of Börgers et al. (2013). Since the two types in the non-congruent case have very different beliefs, it is easy to design experiments that are a close complement to the private information of one type while being of relatively low value to the other. This is much harder in the congruent case, since the two types are very similar.

4 Discussion

Credit bureaus are data collection agencies which gather information related to the credit worthiness of a consumer or business and sell this information to prospective lenders to evaluate the risks of providing loans or credit lines. In the U.S., Experian and Equifax are two major credit bureaus for consumer and business information.[15] The sale of information on debt monitoring by credit bureaus is a natural application considered in the literature (Bergemann and Bonatti, 2019).

Directly relevant to the setting considered in this paper, a credit bureau, the data-seller, offers access to different versions of their database to banks and other financial institutions, the data-buyers. These versions range from complete access to access to a subset of the database, where the latter can be interpreted as partial information. For example, the access to business score indicators range from detailed scoring systems to summary indicators. In exchange for a fee, a credit bureau provides access to a version of their database and condition their pricing only on the information product they are providing. In particular, a credit bureau doesn't personalize their data offerings to a lender as a function of specific realizations regarding credit-worthiness of the lender's prospective client or make them conditional on the information or actions of other lender.[16] Since the database is oftentimes highly sensitive, the credit bureau can typically only provide private access to the database for clients with a "permissable purpose", such as banks or government agencies, which cannot share this information between themselves or with anyone else. These two features directly motivate the model's focus on private signals in a restricted contracting environment.

Furthermore, while an institution could expect that others also access a credit bureau's database, the specific information product that they acquire is typically not observed. Given the sensitive nature of the information, a credit-bureau is unlikely to make specific information acquisitions public. Hence, financial institutions typically remain uncertain about the information observed by others and make inferences about it using their own private information, motivating the assumption that information acquisition decisions are covert.

Financial institutions in turn use the acquired information to decide whether to offer a financial product to their prospective client. In some cases, lenders may be competing for the same client and thus have anti-coordination incentives. In other cases, several lenders may contribute to a large investment project headed by one commercial client or serve different commercial clients who offer complementary goods, in which case they have coordination incentives. Private information that lenders acquire on a client before buying additional

[15] Business credit reports contain information such as ownership, subsidiaries, company finances, vendor payment data, risk scores, and any liens or bankruptcies. Consumer credit reports focus only on an individual's personal credit and includes information regarding loans, credit cards, delinquent accounts, etc.

[16] If the data offered is not verifiable or is costly to verify, it is not incentive compatible for a data-seller to tailor access and prices to specific realizations.

information from a credit bureau is also likely correlated. If they rely on the same information sources and guidelines, then it is likely positively correlated. In contrast, if they systematically differ in their information-gathering, use different private contacts as sources of information, or if their previous experiences with similar clients or projects differ, then private information can be negatively correlated. If all possible private information that can be gathered by these lenders (excluding the information provided by the credit bureau) would lead them to always or never offer a financial product, for example because the project is much above or below the bar for approval, then we can interpret this as lenders having congruent beliefs. Otherwise, more marginal cases can be interpreted as lenders having non-congruent beliefs.

In this context, my model generates a number of insights by considering the role of private information, strategic incentives, and their interaction. First, my results suggest that the credit bureau can price information at the willingness to pay of lenders. Second, access to the full database should be offered and priced at the highest willingness to pay for it. Third, the subset of information offered to the lender with the lower willingness to pay for the full database should be chosen such that the lender with the highest willingness to pay is indifferent between the two information offerings.

When the lenders' private information is such that they would have taken different actions without acquiring additional information, the second item of the menu should offer partial information. The reason is that lenders can disagree in their preferences for information. For instance, two lenders can disagree on the value of different subsets of a database when they observe different characteristics of a borrower's credit history that would lead them to make different choices. Then, databases which include different subsets of a borrower's characteristics would have a different value for each lender. In particular, a database may be more valuable for a lender if it complements their private information by giving them information on characteristics that they had not observed.

When the lenders' private information is such that they would have taken the same actions without acquiring additional information, the second item of the menu should offer partial information if lenders' private information is negatively correlated and they have coordination incentives or if lenders' private information is positively correlated and they have anti-coordination incentives. In the first case, this implies that the credit bureau has more scope for providing partial access to its database when lenders are likely to observe different information about a client's credit worthiness but have incentives to coordinate their actions, for example in the case of a large commercial client. In the second case, partial access should also be provided when lenders are likely to observe similar information about a prospective client's credit history, but compete for them.

My results suggest that credit bureaus should tailor access to their database based on whether the credit information relates to a consumer or a small business versus a large business. Similarly, it may be optimal for credit bureaus to allow different types of lenders to select into differential access to their database. Both of these takeaways follow explicitly from considering the role of private information and strategic incentives that arise when different data-buyers interact in a market. In fact, in most cases, credit bureaus face lenders who compete for clients with private information that partially overlaps. This corresponds to the case with anti-coordination incentives and positively correlated private information. In this case, my results suggest that it is optimal for a credit bureau to offer multiple levels of access to their data set. Consistent with this, credit bureaus in practice do condition the pricing of their information offerings on whether the information relates to consumers or businesses and whether lenders want a one-time access regarding a prospective client or whether they would like to pay a yearly fee. Moreover, they offer a menu of information offerings with at least two levels of access, including full access, suggesting that credit bureaus do allow different lender types to select into different levels of access as implied by my results.

5 Extensions

5.1 N data-buyers

In this section, I extend my results to a setting with $N \geq 2$ data-buyers. Data-buyers are privately informed about the state and attach probability $\theta \in \{\theta_L, \theta_H\}$ to state ω_1. Data-buyers' private information is correlated. In particular, assume that data-buyers' types $\theta_1, ...\theta_N$ are exchangeable random variables with correlation given by

$$\eta_L = \mathbb{P}(\theta_j = \theta_L | \theta_i = \theta_L) \in (0, 1] \text{ and } \eta_H = \mathbb{P}(\theta_j = \theta_H | \theta_i = \theta_H) \in (0, 1]$$

for all j.[17] Let $k \in \{0, 1, ..., N\}$ be the number of high types among the N data-buyers. Denote by ρ_k the probability of observing k high type data-buyers and $N - k$ low types, where $\rho_k \geq 0$ and $\sum_{k=0}^{N} \rho_k = 1$.

In the action stage, ex-post payoffs depend on whether or not data-buyer i matches the state and on whether or not the majority of data-buyers choose the same action as i. In

[17] Exchangeability is a property of the joint distribution of random variables. Exchangeable random variables, though correlated, have equal distributions, i.e., the probability of $\theta_i = \theta_H$ is constant across i.

particular, ex-post payoffs are given by:

$$u_i(a,\omega) = \begin{cases} 1 & if\ a_i = a_\ell,\ \kappa^\ell_{-i} + 1 > \lceil \frac{N}{2} \rceil\ and\ \omega = \omega_\ell \\ c & if\ a_i = a_\ell,\ \kappa^\ell_{-i} + 1 \leq \lceil \frac{N}{2} \rceil\ and\ \omega = \omega_\ell \\ 0 & if\ a_i = a_\ell,\ \kappa^\ell_{-i} + 1 \leq \lceil \frac{N}{2} \rceil\ and\ \omega = \omega_{\ell'} \\ 0 & if\ a_i = a_\ell,\ \kappa^\ell_{-i} + 1 > \lceil \frac{N}{2} \rceil\ and\ \omega = \omega_{\ell'} \end{cases}$$

where κ^ℓ_{-i} is the number of data-buyers $-i$ who choose action $a_\ell \in \{a_1, a_2\}$ and $c > 0$. As in the two data-buyer case, it is an ex-post dominant strategy to select the action that matches the state. Data-buyers are said to have coordination incentives if they prefer to match the majority and anti-coordination incentives otherwise. That is, data-buyers have coordination (anti-coordination) incentives if $c < 1$ ($c > 1$).

Assume that c is sufficiently small such that an symmetric menu is optimal for the data-seller. The data-seller's problem is to select the optimal menu of experiments to maximize her expected profits subject to the data-buyers' participation and incentive-compatibility constraints. That is:

$$\begin{aligned} \max_{(E^m,t^m)_{m=1} \in \{L,H\}} \sum_{k=0}^{N} \rho_k \left((N-k)t^L + k \cdot t^H\right) \text{ subject to } & V_{\mathcal{M}}(E^L,\theta_L) - t^L \geq 0, \\ & V_{\mathcal{M}}(E^H,\theta_H) - t^H \geq 0, \\ & V_{\mathcal{M}}(E^L,\theta_L) - t^L \geq V_{\mathcal{M}}(E^H,\theta_L) - t^H, \\ & V_{\mathcal{M}}(E^H,\theta_H) - t^H \geq V_{\mathcal{M}}(E^L,\theta_H) - t^L. \end{aligned}$$

Value of information. Suppose that all buyers but i purchase the experiment designed for their corresponding type. The number of data-buyers $-i$ who choose action a_1 conditional on the state ω and on the type of data-buyer i, $\kappa^1_{-i}|(\omega,\theta_i)$, is distributed according to a Conway-Maxwell-Binomial distribution,[18] with parameters $N-1$, ν and

$$p_{\omega,\theta_i} = \mathbb{P}(\theta_j = \theta_H|\theta_i)\mathbb{P}(s_j = s_1|\omega,\theta_j) + \mathbb{P}(\theta_j = \theta_L|\theta_i)\mathbb{P}(s_j = s_1|\omega,\theta_j).$$

The parameter ν characterizes the underlying correlation among Bernoulli trials, which captures the correlation among data-buyers' private information. In particular, if $\nu > 1$, the

[18] The Conway-Maxwell-Binomial distribution generalizes the binomial distribution and allows both positive and negative correlation among the exchangeable Bernoulli trials. See Kadane et al. (2016) and Daly and Gaunt (2015) for details. Note that the probability of a data-buyer j choosing action a_1 conditional on the state and θ_i is constant across j.

Bernoulli random variables are negatively correlated. Conversely, when $\nu < 1$, the Bernoulli random variables are positively correlated. Lastly, if $\nu = 1$, the Conway-Maxwell-Binomial distribution simplifies to a Binomial distribution in which Bernoulli trials are independent.

Define Λ_k^θ as the expected gain of choosing the action that matches state ω_k with $k \in \{1, 2\}$ conditional on data-buyer i being type θ. That is:[19]

$$\Lambda_1^\theta = \mathbb{P}\left(\kappa_{-i}^1 + 1 \leq \left\lceil \frac{N}{2} \right\rceil | \omega = \omega_1\right) c + \mathbb{P}\left(\kappa_{-i}^1 + 1 > \left\lceil \frac{N}{2} \right\rceil | \omega = \omega_1\right)$$

and

$$\Lambda_2^\theta = \mathbb{P}\left(\kappa_{-i}^1 \geq N - \left\lceil \frac{N}{2} \right\rceil | \omega = \omega_2\right) c + \mathbb{P}\left(\kappa_{-i}^1 < N - \left\lceil \frac{N}{2} \right\rceil | \omega = \omega_2\right).$$

Data-buyer i's expected gain of acquiring information when she would choose a_2 and a_1 without observing supplemental information are respectively given by:

$$V_2(E^n, \theta | \kappa_{-i}^1) = \theta \pi_1^n \Lambda_1^\theta - (1 - \theta)(1 - \pi_2^n)\Lambda_2^\theta \text{ and } V_1(E^n, \theta | \kappa_{-i}^1) \ = (1 - \theta)\pi_2^n \Lambda_2^\theta - \theta(1 - \pi_1^n)\Lambda_1^\theta.$$

Then, data-buyer i's willingness to pay for experiment E^n is

$$V_{\mathcal{M}}(E^n, \theta) = \begin{cases} \max\{0, V_2(E^n, \theta | \kappa_{-i}^1)\} & if \ \alpha_{i, \iota_{i\theta}=0} = a_2 \\ \max\{0, V_1(E^n, \theta | \kappa_{-i}^1)\} & if \ \alpha_{i, \iota_{i\theta}=0} = a_1. \end{cases}$$

Assumption 2 *Assume that the value of experiment E^n is increasing in its precision.*[20] *That is:*

$$c < 1 \text{ and } c\left(1 + \left\lceil \frac{N}{2} \right\rceil\right) \geq \left\lceil \frac{N}{2} \right\rceil \tag{3}$$

or

$$c > 1 \text{ and } \left(1 + \left\lceil \frac{N}{2} \right\rceil\right) \geq c\left\lceil \frac{N}{2} \right\rceil. \tag{4}$$

[19] Note that $\kappa_{-i}^1 + \kappa_{-i}^2 = N - 1$. Then, $\kappa_{-i}^2 + 1 \leq \lceil \frac{N}{2} \rceil$ is equivalent to $\kappa_{-i}^1 \geq \lceil \frac{N}{2} \rceil$.

[20] These restrictions on the payoff structure are sufficient but not necessary conditions for the value of experiment E^n to be increasing in its precision. See appendix A.2 for details.

Optimal menu. Proposition 1 extends to the case with N data-buyers if willingness to pay for an experiment goes up as its precision increases. Hence, it is optimal for the data-seller to offer the perfectly informative experiment to the high type, whereas the information provided to the low type depends on the coordination incentives and the distribution of data-buyer types. Specifically, it depends on whether or not interim beliefs are congruent. The information provided to the low type is stated in Proposition 5 and 6. The interpretation is analogous as for the two data-buyer case.

Proposition 5 *Assume that data-buyers' beliefs are non-congruent. In an optimal menu, the high type observes perfect information and the low type partial information.*

Proposition 6 *Assume that data-buyers' beliefs are congruent. In an optimal menu, there exists* $\tilde{\theta} \in (0, 1)$ *such that the high type observes perfect information and the low type observes*

i) partial information when data-buyers' private information is negatively correlated, they have coordination incentives and $\theta_L < \tilde{\theta}$.

ii) partial information when data-buyers' private information is positively correlated, they have anti-coordination incentives and $\theta_L < \tilde{\theta}$.

iii) no information, otherwise.

5.2 Correlated signals

In the main section, I assume that the experiments offered by the data-seller generate signals that are conditionally independent given the state. This allows me to focus on the role of the downstream strategic interaction in determining the value of information and the structure of screening menus. It is, however, also natural to study the role of correlation between the signals obtained by different data-buyers. Such correlation may be valuable to data buyers who seek to coordinate their actions not only with the state but also with each other. Correlation can arise naturally as a byproduct of the information provision process. For example, a data-seller may provide identical reports to two data-buyers who requested information of the same level of informativeness. Additionally, correlation can be intentionally introduced as part of the design of the data provider's information offerings.

To address this question, I extend the model as follows: the data-seller continues to offer a menu of experiments to each data-buyer. Data-buyers simultaneously choose their preferred experiment, and it is this choice that determines the marginal distribution of signals given the state. However, in this extension, the correlation between the two signals is determined by a correlation structure denoted as $\psi : \mathcal{M} \times \Omega \mapsto [-1, 1]$ chosen by the data-seller. The correlation can depend on the realized state and the experiment chosen by both data-buyers. Formally, consider the joint signal distribution in state ω_k when data-buyers choose E^m and E^n, respectively. It is given in Table 4.

$s^{m,1}/s^{n,2}$		s_1	s_2
		$\pi_k^{n,2}$	$1-\pi_k^{n,2}$
s_1	$\pi_k^{m,1}$	$\pi_k^{m,1}\pi_k^{n,2}+\psi_k^{m,n}$	$\pi_k^{m,1}\left(1-\pi_k^{n,2}\right)-\psi_k^{m,n}$
s_2	$1-\pi_k^{m,1}$	$\left(1-\pi_k^{m,1}\right)\pi_k^{n,2}-\psi_k^{m,n}$	$\left(1-\pi_k^{m,1}\right)\left(1-\pi_k^{n,2}\right)+\psi_k^{m,n}$

Table 4: Joint signal distribution in state ω_k when data-buyers choose E^m and E^n.

Consider the value of information $V_{(\mathcal{M},\psi)}(E_i^n,\theta_i)$ of experiment E_i^n with menu $\mathcal{M}$ and correlation structure ψ for data-buyer i of type θ_i. It is easy to compute that

$$V_{(\mathcal{M},\psi)}(E_i^n,\theta_i)=V_{\mathcal{M}}(E_i^n,\theta_i)+(1-c)\mathbb{E}\left[\psi_j^{m,n}|\theta_i\right].$$

Therefore, from the perspective of participation alone, the role of correlation is straightforward. Correlated signals are preferred if and only if the data-buyers intend to coordinate their actions in the base game ($c<1$). With private information, correlation also affects incentive compatibility. The following proposition shows that this effect does not overturn the direct impact on the value of information.

Proposition 7 *When the data-seller can offer correlated signals, the optimal symmetric menu is as follows. The high type obtains perfect information. Partial information is offered to the low type if beliefs are non congruent and if beliefs are congruent and*

i) data-buyers have coordination incentives and negatively correlated private information,

ii) data-buyers have anti-coordination incentives and positively correlated private information.

Otherwise, no information is offered to the low type. All signals are maximally correlated when data-buyers have coordination incentives and negatively correlated when they have anti-coordination incentives.

The optimal menu satisfies the same properties as in the case in which the data-seller could only offer experiments with conditionally independent signals across data-buyers, showing that my results are robust to the option to design the correlation between signals. However, the data-seller offers experiments that are maximally (negatively) correlated when data-buyers have coordination (anti-coordination) incentives.

5.3 State-dependent coordination incentives

Assume now that data-buyers' coordination incentives depend on the state. In particular, assume that the payoffs $u : A \times \Omega \to \mathbb{R}$, defined in Table 5, are symmetric and characterized by $c > 0$ and $g > 0$.

$\omega = \omega_1$	a_1	a_2
a_1	$1, 1$	$c, 0$
a_2	$0, c$	$0, 0$

$\omega = \omega_2$	a_1	a_2
a_1	$0, 0$	$0, g$
a_2	$g, 0$	$1, 1$

Table 5: Action stage payoffs.

Under these assumptions, it is an ex-post dominant strategy for each data-buyer to match the state ω. The payoff parameters c and d determine a data-buyer's preference over the action of the other data-buyer. In particular, data-buyer i prefers when j selects the same action (a different action) in state ω_1 when $c < 1$ ($c > 1$). Analogously, data-buyer i prefers when j selects the same action (a different action) in state ω_2 when $d < 1$ ($d > 1$). Formally, data-buyers are said to have coordination (anti-coordination) incentives in state ω_k with $k \in \{1, 2\}$ if the expected gain of choosing an action increases (decreases) in the probability that the other data-buyer chooses the same action. That is, data-buyers have coordination incentives in state ω_1 (ω_2) if $c < 1$ ($d < 1$) and anti-coordination incentives if $c > 1$ ($d < 1$).

Data-buyer i's value of experiment E_i^n when the data-seller offers the menu $\mathcal{M}$ is now given by

$$V_{\mathcal{M}}(E_i^n, \theta_n) = \max \left\{ 0, \sum_{m \in \{L,H\}} \mathbb{P}(\theta_j = \theta_m | \theta_n) v_k(E_i^n, \theta_n; m) \text{ s.t. } k \text{ solves } \alpha_{i,\iota_{i\theta}=0} = a_k \right\},$$

$$v_2(E_i^n, \theta_n; m) = \theta_n \pi_1^{n,i} \left[\pi_1^{m,j} + \left(1 - \pi_1^{m,j}\right) c \right] - (1 - \theta_n) \left(1 - \pi_2^{n,i}\right) \left[\left(1 - \pi_2^{m,j}\right) g + \pi_2^{m,j} \right] \text{ and}$$

$$v_1(E_i^n, \theta_n; m) = (1 - \theta_n) \pi_2^{n,i} \left[\left(1 - \pi_2^{m,j}\right) g + \pi_2^{m,j} \right] - \theta_n \left(1 - \pi_1^{n,i}\right) \left[\pi_1^{m,j} + \left(1 - \pi_1^{m,j}\right) c \right].$$

In this context, it is also optimal for the data-seller to offer concentrated experiments by an analogous argument as in Proposition 1. When data-buyers' beliefs are congruent, the data-seller's problem is identical to the baseline model depending only on $c > 0$ and, therefore, the optimal menu is as characterized in Section 3. When data-buyers' beliefs are non-congruent, it can be shown analogously that the optimal menu is symmetric when $c < \frac{2\rho + \nu}{\rho}$. In this case, the data-seller offers perfect information to the high type and partial information to the low type. Otherwise, the optimal menu is asymmetric and only includes the perfectly informative and the uninformative experiment, as before.

Proposition 8 *A menu with at most two informative experiments is not always optimal.*

5.4 Continuum of types

Consider now the case in which there is a continuum of data-buyer types and the data-seller offers a symmetric menu. Formally, assume that the interim belief θ, which denotes the probability of the state $\omega = \omega_1$, is an element of the unit interval $[0, 1]$ distributed according to the joint distribution $F(\theta)$ and corresponding density $f(\theta)$ with full support. Assume that $F(\theta)$ is exchangeable. Let $F_i(\theta_i)$ be the marginal distribution of θ_i and let $f(\theta_j|\theta_i) = f_{\theta_i}(\theta_j)$ be its density. The value of experiment E^m for data-buyer i of type θ_i is

$$\begin{aligned} V_{\mathcal{M}}(E^m, \theta_i) = {} & \theta_i \pi_1^m \int_0^1 f_{\theta_i}(\theta_j)(c\,(1-\pi_1(\theta_j)) + \pi_1(\theta_j))\mathrm{d}\theta_j + (1-\theta_i)\pi_2^m \int_0^1 f_{\theta_i}(\theta_j)(c\,(1-\pi_2(\theta_j)) + \pi_2(\theta_j))\mathrm{d}\theta_j \\ & - \max\{\theta_i \int_0^1 f_{\theta_i}(\theta_j)(c\,(1-\pi_1(\theta_j)) + \pi_1(\theta_j))\mathrm{d}\theta_j, (1-\theta_i) \int_0^1 f_{\theta_i}(\theta_j)(c\,(1-\pi_2(\theta_j)) + \pi_2(\theta_j))\mathrm{d}\theta_j\} \end{aligned}$$

Let θ_i' be the reported type of data-buyer i and θ_i her realized type. Define

$$q(\theta_i, \theta_i') = \pi_1(\theta_i') \int_0^1 f_{\theta_i}(\theta_j)\,(c\,(1-\pi_1(\theta_j)) + \pi_1(\theta_j))\mathrm{d}\theta_j - \pi_2(\theta_i') \int_0^1 f_{\theta_i}(\theta_j)\,(c\,(1-\pi_2(\theta_j)) + \pi_2(\theta_j))\mathrm{d}\theta_j$$

as the differential informativeness of an experiment E^m. In Bergemann et al. (2018), q is a one-dimensional variable which fully characterizes an experiment given θ_i. Hence, it is possible to rewrite the menu in terms of this variable and derive the optimal menu when the type of a data-buyer is continuous. However, when a data-seller faces multiple data-buyers, this approach doesn't apply unless types are independent since $q(\theta_i, \theta_i')$ depends both on the true type and the reported type through the correlation with the other data-buyer's type. This is because a data-buyer's beliefs about the information observed by others depends on the correlation of data-buyers' information given her true type whereas the information she observes depends on her report. Therefore, we assume independence, which guaranties that $q(\theta_i, \theta_i') = q(\theta_i'', \theta_i') := q(\theta_i')$ for all θ_i and θ_i'', where

$$\begin{aligned} q(\theta_i) \in & \left[-\int_0^1 f(\theta_j)(c\,(1-\pi_2(\theta_j)) + \pi_2(\theta_j))\mathrm{d}\theta_j, \int_0^1 f(\theta_j)(c\,(1-\pi_1(\theta_j)) + \pi_1(\theta_j))\mathrm{d}\theta_j\right] \\ & := [-U_2, U_1]. \end{aligned}$$

Note that both U_1 and U_2 depend on the experiment of others and that the end points of this interval correspond to the uninformative experiment whereas $q(\theta_i) = U_1 - U_2$, to the perfectly informative one. Hence, the value of an experiment $q(\theta_i')$ for data-buyer θ_i can be written as

$$V_{\mathcal{M}}(q(\theta_i'), \theta_i) = \theta_i q(\theta_i') + \pi_2(\theta_i')U_2 - \max\{\theta_i U_1, (1-\theta_i)U_2\}$$

Therefore, there exists $\theta_i^* := \frac{U_2}{U_1+U_2} \in (0,1)$ which corresponds to the type that is ex-ante indifferent between action a_1 and a_2 who assigns the highest value to any experiment q. Moreover, data-buyers who are ex-ante certain about the state, $\theta_i \in \{0,1\}$, assign no value to information. Furthermore, $V_{\mathcal{M}}(q,\theta)$ is linear and increasing (decreasing) in the interval $[0,\theta^*]$ ($[\theta^*,1]$) and exhibits a kink at $\theta = \theta^*$. These properties imply that it is optimal for the data-seller to set $\pi_1(\theta_i) = 1$ for all $\theta_i \geq \theta_i^*$ and $\pi_2(\theta_i) = 1$ for all $\theta_i < \theta_i^*$, where $\theta_i \geq \theta_i^*$ is equivalent to $q(\theta_i) \geq U_1 - U_2$. This is formalized in Lemma 9.

Lemma 9 *There exists $\hat{c} \geq 1$ such that for all $c \leq \hat{c}$, the optimal menu is concentrated and $\pi_1^*(\theta_i) = 1$ for all $q(\theta_i) \geq U_1 - U_2$ and $\pi_2^*(\theta_i) = 1$ for all $q(\theta_i) < U_1 - U_2$. Moreover, $\pi_1^*(\theta_i)$ is non-decreasing and $\pi_2^*(\theta_i)$ is non-increasing.*

In what follows, we restrict attention to $c \leq \hat{c}$. Using Lemma 9, we can write the value of experiment $q(\theta_i')$ for θ_i as follows:

$$V_{\mathcal{M}}(q(\theta_i'), \theta_i) = \theta_i q(\theta_i') + U_2 + \min\{U_1 - U_2 - q(\theta_i'), 0\} - \max\{\theta_i\ U_1, (1-\theta_i)U_2\}.$$

Note that the value of an experiment q satisfies the single-crossing property in (q,θ_i), which implies that higher types assign a higher value to experiments with a higher q. Moreover, experiments with higher q are those which contain a signal that provides stronger evidence of state ω_2, the one deemed less likely given their interim beliefs.

Lemma 2 and Lemma 3 imply that it is without loss of generality to focus on responsive menus. The data-seller's problem is then

$$\max_{q(\theta_i),t(\theta_i)} \int_0^1 t(\theta_i)\mathrm{d}F(\theta_i) \text{ s.t. } V_{\mathcal{M}}\left(q(\theta_i),\theta_i\right) - t(\theta_i) \geq 0 \text{ for all } \theta_i \in [0,1] \text{ and}$$
$$V_{\mathcal{M}}\left(q(\theta_i),\theta_i\right) - t(\theta_i) \geq V\left(q(\theta_i'),\theta_i\right) - t(\theta_i') \text{ for all } \theta_i,\theta_i' \in [0,1].$$

Lemma 10 characterizes the implementable responsive menus, which follows from the envelope theorem.

Lemma 10 *A menu $\{q(\theta_i)\}_{\theta_i\in[0,1]}$ is implementable if and only if $q(\theta_i)$ is non-decreasing and*

$$\int_0^1 q(\theta_i)\mathrm{d}\theta_i = U_1 - U_2 \tag{5}$$

Using Lemma 10, the data-seller's problem can be written as:

$$\max_{q(\theta_i)} \int_0^1 (\theta_i f(\theta_i) + F(\theta_i))q(\theta_i)\mathrm{d}\theta_i + \int_0^1 \min\{0, f(\theta_i)\,[U_1 - U_2 - q(\theta_i)]\}\mathrm{d}\theta_i$$
$$\text{s.t. } q(\theta_i) \in [-U_2, U_1] \text{ and (5).}$$

It is easy to see that it is optimal for the data-seller to offer no information to the types who are certain about the state and full information to $\theta_i = \theta^*$. Hence, we know that an optimal menu includes at least two experiments. In the case with a single data-buyer, Bergemann et al. (2018) show that the data seller uses at most three experiments. Unlike in their case, the objective function in the setting with two buyers is non-linear in q since U_1 and U_2 depend on the experiments offered to the other data-buyer. Hence, their results do not generalize to the a multiple data-buyer setting. In particular, two informative experiments are no longer sufficient to achieve the optimum. In Appendix A.4, I develop an example and establish computationally that the data seller can do better by using more than two informative experiments.

6 Conclusion

This paper considers a setting in which a monopolist data-seller offers supplemental information to privately informed data-buyers. Consistent with previous work, data-buyers' demand for information depends on the precision of their private information. However, it also depends on the correlation of the data-buyers' private information and their strategic interactions.

The data-seller offers a menu of experiments to screen the data-buyers' types. The interaction between coordination incentives and the correlation of private information is the main determinant of the features of the optimal menu. Whenever this interaction increases demand for information for the low type, the data-seller is able to offer partial information to the low type even when beliefs are congruent. In any optimal menu, the data-seller reveals the state to the data-buyer type with the highest willingness to pay for such information. If private information leads data-buyers to choose different actions in the absence of supplemental information, the data-seller can exploit the position of information to provide partial information to the data-buyer with the lowest willingness to pay, without conceding rents to the high type. Indeed, the data-seller can provide partial information to the low type if data-buyers have coordination incentives and their private information is negatively correlated or if data-buyers have anti-coordination incentives and their private information is positively correlated.

These results highlight that the interaction of strategic incentives and correlated private information can relax the incentive compatibility constraints, allowing the data-seller to increase profits by not excluding the low type segment from the market. Considering strategic interactions between data-buyers when designing information offerings is of central importance, both qualitatively and quantitatively, given that data-buyers often interact with others in markets. This non-exclusion result is consistent with results from the multidimensional screening, in which the data-seller offers distorted partial information to the low type designed to ensure that the high type is indifferent between the information offerings.

References

Anat R Admati and Paul Pfleiderer. A monopolistic market for information. *Journal of Economic Theory*, 39(2):400–438, 1986.

Anat R Admati and Paul Pfleiderer. Direct and indirect sale of information. *Econometrica: Journal of the Econometric Society*, pages 901–928, 1990.

Rabah Amir and Natalia Lazzati. Endogenous information acquisition in bayesian games with strategic complementarities. *Journal of Economic Theory*, 163:684–698, 2016.

Pierpaolo Battigalli. Strategic independence and perfect bayesian equilibria. *Journal of Economic Theory*, 70(1):201–234, 1996.

Dirk Bergemann and Alessandro Bonatti. Selling cookies. *American Economic Journal: Microeconomics*, 7(3):259–94, 2015.

Dirk Bergemann and Alessandro Bonatti. Markets for information: An introduction. *Annual Review of Economics*, 11, 2019.

Dirk Bergemann, Alessandro Bonatti, and Alex Smolin. The design and price of information. *American Economic Review*, 108(1):1–48, 2018.

Kostas Bimpikis, Davide Crapis, and Alireza Tahbaz-Salehi. Information sale and competition. *Management Science*, 65(6):2646–2664, 2019.

Alessandro Bonatti, Munther Dahleh, Thibaut Horel, and Amir Nouripour. Selling information in competitive environments. *arXiv preprint arXiv:2202.08780*, 2022.

Alessandro Bonatti, Munther Dahleh, Thibaut Horel, and Amir Nouripour. Coordination via selling information. *arXiv preprint arXiv:2302.12223*, 2023.

Tilman Börgers, Angel Hernando-Veciana, and Daniel Krähmer. When are signals complements or substitutes? *Journal of Economic Theory*, 148(1):165–195, 2013.

Ozan Candogan and Philipp Strack. Optimal disclosure of information to a privately informed receiver. *Available at SSRN 3773326*, 2021.

Jacques Crémer and Richard P McLean. Full extraction of the surplus in bayesian and dominant strategy auctions. *Econometrica: Journal of the Econometric Society*, pages 1247–1257, 1988.

Fraser Daly and Robert E Gaunt. The conway-maxwell-poisson distribution: distributional theory and approximation. *arXiv preprint arXiv:1503.07012*, 2015.

Péter Eső and Balazs Szentes. Optimal information disclosure in auctions and the handicap auction. *The Review of Economic Studies*, 74(3):705–731, 2007.

Christian Hellwig and Laura Veldkamp. Knowing what others know: Coordination motives in information acquisition. *The Review of Economic Studies*, 76(1):223–251, 2009.

Joseph B Kadane et al. Sums of possibly associated bernoulli variables: The conway–maxwell-binomial distribution. *Bayesian Analysis*, 11(2):403–420, 2016.

Jakub Kastl, Marco Pagnozzi, and Salvatore Piccolo. Selling information to competitive firms. *The RAND Journal of Economics*, 49(1):254–282, 2018.

Anton Kolotilin, Tymofiy Mylovanov, Andriy Zapechelnyuk, and Ming Li. Persuasion of a privately informed receiver. *Econometrica*, 85(6):1949–1964, 2017.

Daniel Krähmer. Information disclosure and full surplus extraction in mechanism design. *Journal of Economic Theory*, 187:105020, 2020.

Hao Li and Xianwen Shi. Discriminatory information disclosure. *American Economic Review*, 107(11):3363–85, 2017.

Laurent Mathevet, Jacopo Perego, and Ina Taneva. On information design in games. *Journal of Political Economy*, 128(4):1370–1404, 2020.

David P Myatt and Chris Wallace. Endogenous information acquisition in coordination games. *The Review of Economic Studies*, 79(1):340–374, 2011.

Roger B Myerson. Optimal auction design. *Mathematics of operations research*, 6(1):58–73, 1981.

Roger B Myerson. Optimal coordination mechanisms in generalized principal–agent problems. *Journal of mathematical economics*, 10(1):67–81, 1982.

Roger B Myerson. *Game theory: analysis of conflict*. Harvard university press, 1991.

Carlos Segura-Rodriguez. Selling data. 2021.

Ina Taneva. Information design. *American Economic Journal: Microeconomics*, 11(4):151–85, 2019.

Joel Watson. Perfect bayesian equilibrium: General definitions and illustrations. *Unpublished, University of California at San Diego*, 133, 2016.

Takuro Yamashita and Shuguang Zhu. Type-contingent information disclosure. 2021.

Kai Hao Yang. Selling consumer data for profit: Optimal market-segmentation design and its consequences. *American Economic Review*, 112(4):1364–93, 2022.

Ming Yang. Coordination with flexible information acquisition. *Journal of Economic Theory*, 158:721–738, 2015.

Shuguang Zhu. Private disclosure with multiple agents. *Working paper*, 2021.

A Appendix

A.1 Proofs with two data-buyers

Proof. Lemma 1. Define γ_{ki} as the probability that data-buyer type θ_i assigns to data-buyer j selecting action a_1 conditional on the state being ω_k. That is,

$$\gamma_{ki} = \mathbb{P}_i(a_j = a_1 | \omega = \omega_k).$$

The value of experiment $\overline{E}$ for data-buyer type θ_i is

$$\begin{aligned} V_{\mathcal{M}}(\overline{E}, \theta_i) &= \min\{(1-\theta_i)[c\gamma_{2i} + (1-\gamma_{2i})], \theta_i[\gamma_{1i} + (1-\gamma_{1i})c]\} \\ &= \begin{cases} \theta_i[\gamma_{1i} + (1-\gamma_{1i})c] & if\ \theta_i < \hat{\theta}(c, \theta_j) \\ (1-\theta_i)[c\gamma_{2i} + (1-\gamma_{2i})] & if\ \theta_i \geq \hat{\theta}(c, \theta_j) \end{cases} \end{aligned}$$

where $\hat{\theta}(c, \theta_j)$ is the data-buyer type θ_i that is initially indifferent between selecting action a_1 and a_2. First, if $\theta_2 \geq \hat{\theta}_2(c, \theta_1)$, $V_{\mathcal{M}}(\overline{E}, \theta_2) \geq V_{\mathcal{M}}(\overline{E}, \theta_1)$ if and only if

$$\begin{aligned} &i)\ \theta_2 \leq 1 - \frac{\gamma_{11} + (1-\gamma_{11})c}{\gamma_{12} + (1-\gamma_{12})c}\theta_1 \text{ when } \theta_1 < \hat{\theta}(c, \theta_2) \text{ and} \\ &ii)\ \theta_2 \leq 1 - \frac{c\gamma_{21} + (1-\gamma_{21})}{\gamma_{12} + (1-\gamma_{12})c}(1-\theta_1) \text{ when } \theta_1 \geq \hat{\theta}(c, \theta_2). \end{aligned}$$

In this case, $\theta_2 = \theta_H$ and $\theta_1 = \theta_L$. Second, if $\theta_2 < \hat{\theta}(c, \theta_1)$, $V_{\mathcal{M}}(\overline{E}, \theta_2) \geq V_{\mathcal{M}}(\overline{E}, \theta_1)$ if and only if

$$\begin{aligned} &i)\ \theta_2 \geq \frac{\gamma_{11} + (1-\gamma_{11})c}{c\gamma_{22} + (1-\gamma_{22})}\theta_1 \text{ when } \theta_1 < \hat{\theta}(c, \theta_2) \text{ and} \\ &ii)\ \theta_2 \geq \frac{c\gamma_{21} + (1-\gamma_{21})}{c\gamma_{22} + (1-\gamma_{22})}(1-\theta_1) \text{ when } \theta \geq \hat{\theta}(c, \theta_2). \end{aligned}$$

■

Proof. Lemma 2. I show that for any menu $\hat{\mathcal{M}}$ and BNE $\hat{\sigma}$, there exists a direct menu $\mathcal{M}^D$ and σ^D such that (i) every data-buyer i of type $\theta \in \{\theta_L, \theta_H\}$ purchases the experiment designed for her type; (ii) for every type vector, the distribution over outcomes under $\hat{\mathcal{M}}$ if $\hat{\sigma}$ is played is the same as the distribution over outcomes that results under $\mathcal{M}^D$ and σ^D.

Suppose instead that $M \geq 3$. First, suppose that $\hat{\iota}_{i\theta} \in \{0, 1, ..., M\}$ for all i type θ. Hence, up to two elements of the menu per data-buyer are traded in equilibrium. Then, it is possible to eliminate the redundant elements of the menu and offer menu $\mathcal{M}^D$ which only includes the ones that are purchased in equilibrium. In this case, it is trivial that the distribution over outcomes remains unchanged. Second, assume that there exist i type θ such that $\hat{\iota}_i \in \Delta(\{0,1,..,M\})$. Construct an alternative menu $\mathcal{M}^D$ in which the experiments that are chosen by i type θ with positive probability are replace by one experiment that randomizes over those experiments such that induces the same distribution over outcomes. That is, define

$$\pi^D((s_i,s_j)|\omega,(\iota_i,\iota_j)) = \sum_{m=0}^{M} \mathbb{P}(\hat{\iota}_i = m|\omega)\hat{\pi}((s_i,s_j)|\omega,(m,\iota_j)).$$

Note that the overall distribution over outcomes remains unchanged. Thus, $\mathcal{M}^D$ implements the same outcome as $\hat{\mathcal{M}}$ and, since $\hat{\sigma}$ is a Bayes Nash equilibrium, σ^D is also an equilibrium. Therefore, it is without loss of generality to consider menus with at most two elements per data-buyer. ■

Proof. Lemma 3. Consider i type θ and an experiment $E_i^m \in \mathcal{M}_i$ where the menu $\mathcal{M} = (\mathcal{M}_1, \mathcal{M}_2)$ is a incentive-compatible and individually rational. Let i type θ choose a single action after each signal. Given the equilibrium strategies, let $S_k^{m,i}$ denote the subset of signals in experiment E_i^m that induces buyer i type θ to choose action $a_k \in \{a_1, a_2\}$ where $\cup_{k=1}^2 S_k^{m,i} = S^{m,i}$. Construct $\widehat{E}_i^m = (\widehat{S}^{m,i}, \widehat{\pi}^{m,i})$ where $\widehat{S}^{m,i} = \{s_1, s_2\}$ and

$$\widehat{\pi}^{m,i}(s_\ell|\omega) = \int_{S_k^{m,i}} \pi^{m,i}(s|\omega)\mathrm{d}s \text{ for all } s_\ell \in \{s_1,s_2\} \text{ and } \omega \in \{\omega_1,\omega_2\}.$$

E_i^m and $\hat{E}_i^m$ are constructed such that both experiments induce the same outcome for buyer i type θ. Thus, i attaches the same value to both experiments, i.e., $V_{\mathcal{M}}(E_i^m,\theta) = V_{\hat{\mathcal{M}}}(\widehat{E}_i^m,\theta)$. Furthermore, $\widehat{E}_i^m$ is a weakly less informative than E_i^m and Blackwell's theorem implies that $V_{\mathcal{M}}(E_i^m,\theta') \leq V_{\hat{\mathcal{M}}}(\widehat{E}_i^m,\theta')$ for all θ'. This relaxes the incentive constraints of types $\theta' \neq \theta$. Therefore, for any $\mathcal{M}$, it is possible to construct $\widehat{\mathcal{M}}$ that replaces E_i^m with $\widehat{E}_i^m$ that is also incentive compatible and individually rational and yields weakly larger profits. ■

Proof. Lemma 4. A data-buyer i of type θ_H chooses action a_1 without supplemental information when her expected payoffs of choosing a_1 exceeds the expected payoffs of choosing a_2. That is,

$$\begin{aligned}(1-\theta_H)&\left[\frac{\nu}{\nu+\rho}\left(\left(1-{\pi_2}^{L,2}\right)c+{\pi_2}^{L,2}\right)+\frac{\rho}{\nu+\rho}\left(\left(1-{\pi_2}^{H,2}\right)c+{\pi_2}^{H,2}\right)\right]\\ &<\theta_H\left[\frac{\nu}{\nu+\rho}\left(\left(1-{\pi_1}^{L,2}\right)c+{\pi_1}^{L,2}\right)+\frac{\rho}{\nu+\rho}\left(\left(1-{\pi_1}^{H,2}\right)c+{\pi_1}^{H,2}\right)\right].\end{aligned}$$

Note that

$$(1-\theta_H)\left[\frac{\nu}{\nu+\rho}\left(\left(1-{\pi_2}^{L,2}\right)c+{\pi_2}^{L,2}\right)+\frac{\rho}{\nu+\rho}\left(\left(1-{\pi_2}^{H,2}\right)c+{\pi_2}^{H,2}\right)\right]>(1-\theta_H)\min\{c,1\}$$

$$\text{and } \theta_H\left[\frac{\nu}{\nu+\rho}\left(\left(1-{\pi_1}^{L,2}\right)c+{\pi_1}^{L,2}\right)+\frac{\rho}{\nu+\rho}\left(\left(1-{\pi_1}^{H,2}\right)c+{\pi_1}^{H,2}\right)\right]<\theta_H\max\{c,1\}.$$

Therefore, a necessary condition for the high type to choose action a_1 without supplemental information is $\theta_H \geq \min\left\{\frac{1}{1+c}, \frac{c}{1+c}\right\}$. Analogously, we can conclude that the low type must also satisfy $\theta_L \geq \min\left\{\frac{1}{1+c}, \frac{c}{1+c}\right\}$ to select the same action and $\theta_L \leq \max\left\{\frac{1}{1+c}, \frac{c}{1+c}\right\}$ to select a_2.

Moreover, when both types choose action a_1 without supplemental information, the definition of the high type requires that

$$(1-\theta_H)\left[\frac{\nu}{\nu+\rho}\left(\left(1-{\pi_2}^{L,2}\right)c+{\pi_2}^{L,2}\right)+\frac{\rho}{\nu+\rho}\left(\left(1-{\pi_2}^{H,2}\right)c+{\pi_2}^{H,2}\right)\right]$$
$$>(1-\theta_L)\left[\frac{1-2\nu-\rho}{1-\nu-\rho}\left(\left(1-{\pi_2}^{L,2}\right)c+{\pi_2}^{L,2}\right)+\frac{\nu}{1-\nu-\rho}\left(\left(1-{\pi_2}^{H,2}\right)c+{\pi_2}^{H,2}\right)\right].$$

Then, a necessary condition for the previous expression is

$$(1-\theta_H)\max\{c,1\}>(1-\theta_L)\min\{c,1\},$$

which holds for all $c>0$ if and only if $\theta_H<\theta_L$. Analogously, when the low type chooses action a_2 without supplemental information, $(1-\theta_H)\max\{c,1\}>\theta_L\min\{c,1\}$, which holds for all $c>0$ if and only if $\theta_H<1-\theta_L$. ■

Proof. Lemma 5. Let $\Pi(\mathcal{M})$ denote the seller's expected profits from a menu $\mathcal{M}$ that satisfies all participation and incentive-compatibility constraints. Assume without loss of generality that the high type selects action a_1 without supplemental information. Consider first the case in which the low type also chooses a_1 under the same circumstances. Note that

$$\Pi(\mathcal{M})=\begin{cases}2(1-\theta_L) & if\ \mathcal{M}=\underline{\mathcal{M}}^1\\ 2(1-\theta_H)(\nu\cdot c+\rho) & if\ \mathcal{M}=\underline{\mathcal{M}}^2\\ c(1-\theta_L) & if\ \mathcal{M}=\underline{\mathcal{M}}^3\\ c(1-\theta_H)(\nu+\rho) & if\ \mathcal{M}=\underline{\mathcal{M}}^4\end{cases}$$

When $c<2$, $\Pi(\underline{\mathcal{M}}^1)>\Pi(\underline{\mathcal{M}}^3)$ and $\Pi(\underline{\mathcal{M}}^2)>\Pi(\underline{\mathcal{M}}^4)$. Let $\mathcal{M}^*$ be the data-seller's optimal menu. Then,

$$\Pi(\mathcal{M}^*)\geq\Pi(\underline{\mathcal{M}}^1)\text{ when }\frac{1-\theta_L}{1-\theta_H}\geq\nu\cdot c+\rho\text{ and}$$
$$\Pi(\mathcal{M}^*)\geq\Pi(\underline{\mathcal{M}}^2),\ \text{otherwise.}$$

When $c \geq 2$, $\Pi(\underline{\mathcal{M}}^1) \leq \Pi(\underline{\mathcal{M}}^3)$. Then, $\Pi(\mathcal{M}^*) \geq \Pi(\underline{\mathcal{M}}^3)$ when

$$\frac{1-\theta_L}{1-\theta_H} > \max\left\{\nu+\rho, \frac{2}{c}(\nu \cdot c+\rho)\right\}.$$

Otherwise,

$$\begin{aligned}
&\Pi(\mathcal{M}^*) \geq \Pi(\underline{\mathcal{M}}^2) \text{ when } \nu+\rho \leq \frac{2}{c}(\nu \cdot c+\rho) \text{ and} \\
&\Pi(\mathcal{M}^*) \geq \Pi(\underline{\mathcal{M}}^4), \text{ otherwise.}
\end{aligned}$$

Consider now the case in which the low type chooses a_2 under the same circumstances. Note that

$$\Pi(\mathcal{M}) = \begin{cases} 2\theta_L & if\ \mathcal{M} = \underline{\mathcal{M}}^1 \\ 2(1-\theta_H)(\nu+\rho) & if\ \mathcal{M} = \underline{\mathcal{M}}^2 \\ \theta_L\left[\frac{(1-2\nu-\rho)}{1-\nu-\rho}c + \frac{\nu}{1-\nu-\rho}\right] & if\ \mathcal{M} = \underline{\mathcal{M}}^3 \\ (1-\theta_H)(\nu+\rho \cdot c) & if\ \mathcal{M} = \underline{\mathcal{M}}^4 \end{cases}$$

When $c < 2$, $\Pi(\underline{\mathcal{M}}^1) > \Pi(\underline{\mathcal{M}}^3)$ and $\Pi(\underline{\mathcal{M}}^2) > \Pi(\underline{\mathcal{M}}^4)$. Then,

$$\begin{aligned}
&\Pi(\mathcal{M}^*) \geq \Pi(\underline{\mathcal{M}}^1) \text{ when } \frac{\theta_L}{1-\theta_H} \geq \nu+\rho \text{ and} \\
&\Pi(\mathcal{M}^*) \geq \Pi(\underline{\mathcal{M}}^2), \text{ otherwise.}
\end{aligned}$$

When $c \geq 2$, $\Pi(\underline{\mathcal{M}}^1) \leq \Pi(\underline{\mathcal{M}}^3)$. Then, $\Pi(\mathcal{M}^*) \geq \Pi(\underline{\mathcal{M}}^3)$ when

$$\frac{\theta_L}{1-\theta_H} > \frac{(1-\nu-\rho)}{c(1-2\nu-\rho)+\nu} \max\left\{(c\rho+\nu), (2(\nu+\rho))\right\}.$$

Otherwise,

$$\begin{aligned}
&\Pi(\mathcal{M}^*) \geq \Pi(\underline{\mathcal{M}}^2) \text{ when } \frac{1}{2}(\nu+\rho \cdot c) \leq \nu+\rho \text{ and} \\
&\Pi(\mathcal{M}^*) \geq \Pi(\underline{\mathcal{M}}^4), \text{ otherwise.}
\end{aligned}$$

■

Proof. Lemma 6. The data-seller can always offer a menu $\mathcal{M}_*$ which offers only the perfectly informative experiment to both data-buyers at a price p_* equal to the willingness to pay of the low type, where $p = 1-\theta_L$ when the low type picks action a_1 without additional information and $p_* = \theta_L$, otherwise. Hence, the data-seller can guarantee herself profits of at least $2p_*$. Similarly, the data-seller can always offer a menu $\mathcal{M}_{**}$ which offers only the perfectly informative experiment to one data-buyer at a price p_{**} equal to the willingness to pay of the low type and no information to the other data-buyer, where $p_{**} = (1-\theta_L)c$ when the low type picks action a_1 without additional information and

$$p_{**} = \theta_L \left[\frac{(1-2\nu-\rho)}{1-\nu-\rho} c + \frac{\nu}{1-\nu-\rho} \right],$$

otherwise. Hence, the data-seller can guarantee herself profits of at least p_{**}. Therefore, the data-seller's profits for any menu $\mathcal{M}$ satisfies $\Pi(\mathcal{M}) \geq \widehat{\Pi}$ where

$$\widehat{\Pi} = \begin{cases} (1-\theta_L)\max\{2, c\} & if\ a_1\ is\ chosen\ without\ supplemental\ information \\ \theta_L \max\{2, \left[\frac{(1-2\nu-\rho)}{1-\nu-\rho}c + \frac{\nu}{1-\nu-\rho}\right]\} & if\ a_2\ is\ chosen\ without\ supplemental\ information \end{cases}$$

This implies that it is optimal for the data-seller to offer a menu $\mathcal{M}$ with two distinct items to at least one data-buyer i when her expected profits, given by $\Pi(\mathcal{M}) = (1-\nu-\rho)t_i^L + (\nu+\rho)t_i^H$ is at least $\widehat{\Pi}$ where

$$t_i^L \leq \min\left\{ V_{\mathcal{M}}\left(E_i^L, \theta_L\right), V_{\mathcal{M}}\left(E_i^L, \theta_L\right) - \left(V_{\mathcal{M}}\left(E_i^H, \theta_L\right) - t_i^H\right) \right\} \text{ and}$$
$$t_i^H \leq \min\left\{ V_{\mathcal{M}}\left(E_i^H, \theta_H\right), V_{\mathcal{M}}\left(E_i^H, \theta_H\right) - \left(V_{\mathcal{M}}\left(E_i^L, \theta_H\right) - t_i^L\right) \right\}.$$

■

Proof. Proposition 1. 1. The data-seller's problem is given by

$$\max_{(E_i^m, t_i^m)_{(m,i)\in\{L,H\}\times\{1,2\}}} (1-\nu-\rho)(t_1^L + t_2^L) + (\rho+\nu)(t_1^H + t_2^H)$$

subject to the participation constraints

$$IR_{L,i} : V_{\mathcal{M}}(E_i^L, \theta_L) - t_i^L \geq 0,\ IR_{H,i} : V_{\mathcal{M}}(E_i^H, \theta_H) - t_i^H \geq 0$$

and the incentive-compatibility constraints

$$IC_{L,i} : V_{\mathcal{M}}(E_i^L, \theta_L) - t_i^L \geq V_{\mathcal{M}}(E_i^H, \theta_L) - t_i^H$$
$$IC_{H,i} : V_{\mathcal{M}}(E_i^H, \theta_H) - t_i^H \geq V_{\mathcal{M}}(E_i^L, \theta_H) - t_i^L$$

for all $i \in \{1,2\}$. Assume without loss of generality that θ_H selects action a_1 without supplemental information.

Consider first the case in which the low type θ_L also picks action a_1 under the same circumstance. Assume that in the optimal menu $\mathcal{M}^*$, $\pi_1^{L,i} < 1$ and consider another menu $\mathcal{M}$' which replaces $\pi_1^{L,i}$ with 1. Note that for data-buyer i,

$$\begin{aligned} V_{\mathcal{M}'}\left(E_i^L, \theta_L\right) &= (1-\theta_L)\pi_2^{L,i}\left[\frac{1-2\nu-\rho}{1-\nu-\rho}((1-\pi_2^{L,j})c+\pi_2^{L,j}) + \frac{\nu}{1-\nu-\rho}((1-\pi_2^{H,j})c+\pi_2^{H,j})\right] \\ &\geq V_{\mathcal{M}^*}\left(E_i^L, \theta_L\right), \\ V_{\mathcal{M}'}\left(E_i^L, \theta_H\right) &= (1-\theta_H)\pi_2^{L,i}\left[\frac{\nu}{\nu+\rho}((1-{\pi_2}^{L,j})c+{\pi_2}^{L,j}) + \frac{\rho}{\nu+\rho}((1-{\pi_2}^{H,j})c+{\pi_2}^{H,j})\right] \\ &\geq V_{\mathcal{M}^*}\left(E_i^L, \theta_H\right), \end{aligned}$$

$V_{\mathcal{M}'}\left(E_i^H, \theta_L\right) = V_{\mathcal{M}^*}\left(E_i^H, \theta_L\right)$ and $V_{\mathcal{M}'}\left(E_i^H, \theta_H\right) = V_{\mathcal{M}^*}\left(E_i^H, \theta_H\right)$. The optimality of $\mathcal{M}^*$ implies that it satisfies the participation and incentive-compatibility constraints, which, in turn imply that

$$\begin{aligned} V_{\mathcal{M}'}\left(E_i^L, \theta_L\right) \geq V_{\mathcal{M}^*}\left(E_i^L, \theta_L\right) \geq t_i^L \\ V_{\mathcal{M}'}\left(E_i^H, \theta_H\right) = V_{\mathcal{M}^*}\left(E_i^H, \theta_H\right) \geq t_i^H \end{aligned}$$

and

$$\begin{aligned} V_{\mathcal{M}'}\left(E_i^L, \theta_L\right) - t_i^L \geq V_{\mathcal{M}^*}\left(E_i^L, \theta_L\right) - t_i^L \geq V_{\mathcal{M}^*}\left(E_i^H, \theta_L\right) - t_i^H = V_{\mathcal{M}'}\left(E_i^H, \theta_L\right) - t_i^H \\ V_{\mathcal{M}'}\left(E_i^H, \theta_H\right) - t_i^H = V_{\mathcal{M}^*}\left(E_i^H, \theta_H\right) - t_i^H \geq V_{\mathcal{M}^*}\left(E_i^L, \theta_H\right) - t_i^L \end{aligned}$$

Hence, replacing $\pi_1^{L,i}$ with 1 relaxes participation and incentive-compatibility constraint of the high type and has no effect in the participation constraint of the high type. Then, it is possible for the data-seller to increase t_i^L by $\epsilon > 0$ sufficiently small such that all constraints satisfied. Then, the data-seller obtains higher profits from data-buyer i with menu $\mathcal{M}'$ than $\mathcal{M}^*$.

The effect of this change on data-buyer j's willingness to pay for information depends on c. In particular, $V_{\mathcal{M}^*}\left(E_j^m, \theta_k\right) \geq V_{\mathcal{M}'}\left(E_j^m, \theta_k\right)$ for all $(m,k) \in \{L,H\}^2$ if and only if $c < 1$. Hence, the data-seller must decrease transfers from data-buyer j when $c < 1$ and increase them otherwise. It is trivial then that $\mathcal{M}'$ yields higher profits to the data-seller than $\mathcal{M}^*$ when $c < 1$, contradicting the optimality of $\mathcal{M}^*$. When $c > 1$, the data-seller obtains higher profits from data-buyer i and lower from j under menu $\mathcal{M}'$. However, this overall increases the data-seller's profits for all cases in which two distinct experiments are optimal, as derived in Lemma 6.

The proof for the case in which the low type chooses a_2 without supplemental information is analogous.

2. First, I show that if $\overline{E}$ is part of the optimal menu, it is offered to the high type of data-buyer i for some $i \in \{1,2\}$. Suppose instead that only the low type of data-buyer i purchases this experiment, implying that its price cannot exceed $V_{\mathcal{M}}(\overline{E}, \theta_L)$. If the high type does not purchase this experiment, incentive-compatibility implies that $t_i^H < V_{\mathcal{M}}(\overline{E}, \theta_L)$. Thus, only offering $\overline{E}$ at price $V_{\mathcal{M}}(\overline{E}, \theta_L)$ to at least one data-buyer improves the seller's profits. This yields a contradiction.

Second, I show that $\overline{E}$ is part of the optimal menu. Assume without loss of generality that θ_H chooses action a_1 in the absence of supplemental information. Consider first the case in which θ_L also chooses action a_1 under the same circumstances. Proposition 1.1 implies that $\pi_1^{H,i} = \pi_1^{L,i} = 1$ for all $i \in \{1,2\}$. Suppose that in the optimal menu $\mathcal{M}^*$, $\pi_2^{H,i} < 1$ and consider another menu $\mathcal{M}$' which replaces $\pi_2^{H,i}$ with 1 for some $i \in \{1,2\}$. Note that for data-buyer i,

$$
\begin{aligned}
V_{\mathcal{M}'}\left(E_i^L, \theta_L\right) &= V_{\mathcal{M}^*}\left(E_i^L, \theta_L\right) \geq t_i^L \\
V_{\mathcal{M}'}\left(E_i^L, \theta_L\right) - t_i^L &= V_{\mathcal{M}^*}\left(E_i^L, \theta_L\right) - t_i^L \geq V_{\mathcal{M}^*}\left(E_i^H, \theta_L\right) - t_i^H \\
V_{\mathcal{M}'}\left(E_i^H, \theta_H\right) &> V_{\mathcal{M}^*}\left(E_i^H, \theta_H\right) \geq t_i^H \text{ and} \\
V_{\mathcal{M}'}\left(E_i^H, \theta_H\right) - t_i^H &= V_{\mathcal{M}^*}\left(E_i^H, \theta_H\right) - t_i^H \geq V_{\mathcal{M}^*}\left(E_i^L, \theta_H\right) - t_i^L = V_{\mathcal{M}'}\left(E_i^L, \theta_H\right) - t_i^L.
\end{aligned}
$$

Then, increasing $\pi_2^{H,i}$ to 1 increases data-buyer i type θ_H's willingness to pay for experiment E_i^H and relaxes her participation and incentive-compatibility constraints, while having no effect on the participation constraint of the low type. As a result, the data-seller can increase t_i^H appropriately such as the incentive-compatibility constraint of the low type is still satisfied.

To see whether or not this is optimal for the data-seller, it is necessary to analyze the effect of the change on data-buyer j's willingness to pay which depends on c. In particular, if $c < 1$, we have that increasing $\pi_2^{H,i}$ to 1 increases data-buyer j's willingness to pay for information and relaxes her participation constraints because

$$
\begin{aligned}
V_{\mathcal{M}'}\left(E_j^L, \theta_L\right) &> V_{\mathcal{M}^*}\left(E_j^L, \theta_L\right) \geq t_j^L \text{ and} \\
V_{\mathcal{M}'}\left(E_j^H, \theta_H\right) &> V_{\mathcal{M}^*}\left(E_j^H, \theta_H\right) \geq t_j^H
\end{aligned}
$$

Moreover, since

$$
\begin{aligned}
0 \leq V_{\mathcal{M}'}\left(E_j^L, \theta_L\right) - V_{\mathcal{M}^*}\left(E_j^L, \theta_L\right) &\leq V_{\mathcal{M}'}\left(E_j^H, \theta_L\right) - V_{\mathcal{M}^*}\left(E_j^H, \theta_L\right) \text{ and} \\
V_{\mathcal{M}'}\left(E_j^H, \theta_H\right) - V_{\mathcal{M}^*}\left(E_j^H, \theta_H\right) &\geq V_{\mathcal{M}'}\left(E_j^L, \theta_H\right) - V_{\mathcal{M}^*}\left(E_j^L, \theta_H\right) \geq 0,
\end{aligned}
$$

we have that

$$V_{\mathcal{M}^*}\left(E_j^H,\theta_L\right)-V_{\mathcal{M}^*}\left(E_j^L,\theta_L\right)\leq V_{\mathcal{M}'}\left(E_j^H,\theta_L\right)-V_{\mathcal{M}'}\left(E_j^L,\theta_L\right) \text{ and}$$
$$V_{\mathcal{M}'}\left(E_j^H,\theta_H\right)-V_{\mathcal{M}'}\left(E_j^L,\theta_H\right)\geq V_{\mathcal{M}^*}\left(E_j^H,\theta_H\right)-V_{\mathcal{M}^*}\left(E_j^L,\theta_H\right)\geq t_j^H-t_j^L.$$

Hence, when $c<1$, the incentive-compatibility constraints also hold if

$$V_{\mathcal{M}'}\left(E_j^H,\theta_L\right)-V_{\mathcal{M}'}\left(E_j^L,\theta_L\right)\leq t_j^H-t_j^L$$

or if t_j^H is increased appropriately. This implies that $\mathcal{M}'$ satisfies also satisfies data-buyer j's constraints while yielding a higher profit for the data-seller, contradicting the optimality of $\mathcal{M}^*$.

When $c>1$, increasing $\pi_2^{H,i}$ to 1 decreases data-buyer j's willingness to pay for information. The expected change in data-buyer j's willingness to pay is

$$(1-\pi_2^{H,1})(1-c)[(1-\theta_L)\pi_2^{L,2}\nu+(1-\theta_H)\pi_2^{H,2}\rho],$$

whereas the expected change for data-buyer i's is

$$(1-\theta_H)(1-\pi_2^{H,1})[\nu((1-\pi_2^{L,2})c+\pi_2^{L,2})+\rho((1-\pi_2^{H,2})c+\pi_2^{H,2})].$$

Note that the expected gain in willingness to pay for data-buyer i exceeds the expected loss from j whenever

$$c\leq 1+\frac{(1-\theta_H)(\nu+\rho)}{(1-\theta_H)\rho+(1-\theta_L)\nu}.$$

In this case, the data-seller benefits from increasing $\pi_2^{H,i}$ to 1. Otherwise, that is, when c is sufficiently large, the data-seller selects $\mathcal{M}\in\{\underline{\mathcal{M}}^2,\underline{\mathcal{M}}^3,\underline{\mathcal{M}}^4\}$, which offer full information to at least one data-buyer.

The proof is analogous in the case in which the low type chooses a_2 in the absence of supplemental information.

3. First, I show that participation constraint of data-buyers of type θ_L bind. Suppose not, i.e., assume that $t_1^L<V_{\mathcal{M}}(E_1^L,\theta_L)$. Then, the participation constraint of data-buyer 1 of type high must bind, which implies that $t_1^H=V_{\mathcal{M}}(E_1^H,\theta_H)$. Moreover, the incentive compatibility constraints imply

$$IC_{L,1}: V_{\mathcal{M}}(E_1^L,\theta_L)-t_1^L\geq V_{\mathcal{M}}(E_1^H,\theta_L)-t_1^H \text{ and } IC_{H,1}:\ 0\geq V_{\mathcal{M}}(E_1^L,\theta_H)-t_1^L.$$

Note that $V_{\mathcal{M}}(E_1^H, \theta_H)$ cannot be smaller than $V_{\mathcal{M}}(E_1^H, \theta_L)$, since this violates the definition of the high type.[21] Then, since $V_{\mathcal{M}}(E_1^H, \theta_H) > V_{\mathcal{M}}(E_1^H, \theta_L)$, then $V_{\mathcal{M}}(E_1^L, \theta_L) - t_1^L > 0$ and $V_{\mathcal{M}}(E_1^H, \theta_L) - t_1^H < 0$. Hence, it is possible to increase t_1^L by a small $\epsilon > 0$ without violating any compatibility constraint for data-buyer 1 and without inducing any change for the menu offered to data-buyer 2, yielding a contradiction.

Second, I show that the participation constraint of the high type also binds. Suppose not, i.e., $t_1^H > V_{\mathcal{M}}(E_1^H, \theta_H)$. Since the $IR_{L,1}$ binds, incentive-compatibility constraints are given by

$$IC_{L,1} : 0 \geq V_{\mathcal{M}}(E_1^H, \theta_L) - t_1^H \text{ and } IC_{H,1} : V_{\mathcal{M}}(E_1^H, \theta_H) - t_1^H \geq V_{\mathcal{M}}(E_1^L, \theta_H) - V_{\mathcal{M}}(E_1^L, \theta_L).$$

When $V_{\mathcal{M}}(E_1^L, \theta_H) < V_{\mathcal{M}}(E_1^L, \theta_L)$, $V_{\mathcal{M}}(E_1^H, \theta_H) - t_1^H > 0$ and $V_{\mathcal{M}}(E_1^L, \theta_H) - V_{\mathcal{M}}(E_1^L, \theta_L) < 0$. Hence, it is possible to increase t_1^H by a small $\epsilon > 0$ without violating any compatibility constraint for data-buyer 1 and without inducing any change for the menu offered to data-buyer 2, yielding a contradiction.

When $V_{\mathcal{M}}(E_1^L, \theta_H) \geq V_{\mathcal{M}}(E_1^L, \theta_L)$,

$$V_{\mathcal{M}}(E_1^H, \theta_L) \leq t_1^H \leq V_{\mathcal{M}}(E_1^H, \theta_H) - [V_{\mathcal{M}}(E_1^L, \theta_H) - V_{\mathcal{M}}(E_1^L, \theta_L)]$$

and it is optimal for the data-seller to set

$$t_1^H = V_{\mathcal{M}}(E_1^H, \theta_H) - [V_{\mathcal{M}}(E_1^L, \theta_H) - V_{\mathcal{M}}(E_1^L, \theta_L)].$$

[21] Let $\delta_{k,i}$ for $k \in \{1, 2\}$ and $i \in \{L, H\}$ be defined as follows:

$$\delta_{k,L} = \frac{1 - 2\nu - \rho}{1 - \nu - \rho} \left(\left(1 - {\pi_k}^{L,2}\right) c + {\pi_k}^{L,2} \right) + \frac{\nu}{1 - \nu - \rho} \left(\left(1 - {\pi_k}^{H,2}\right) c + {\pi_k}^{H,2} \right) \text{ and}$$
$$\delta_{k,H} = \frac{\nu}{\nu + \rho} \left(\left(1 - {\pi_k}^{L,2}\right) c + {\pi_k}^{L,2} \right) + \frac{\rho}{\nu + \rho} \left(\left(1 - {\pi_k}^{H,2}\right) c + {\pi_k}^{H,2} \right).$$

Consider first the case in which both types of data-buyers choose the same action a_1 without additional information. Then, by definition, $V_{\mathcal{M}}(E_1^H, \theta_H) \leq V_{\mathcal{M}}(E_1^H, \theta_L)$ if and only if

$$[(1 - \theta_L)\delta_{2,L} - (1 - \theta_H)\delta_{2,H}]\, {\pi_2}^{H,1} \geq (\theta_L \delta_{1,L} - \theta_H \delta_{1,H}) \left(1 - {\pi_1}^{H,1}\right)$$

First, if $\theta_L \delta_{1,L} \geq \theta_H \delta_{1,H}$, the previous condition implies that $(1 - \theta_L)\delta_{2,L} \geq (1 - \theta_H)\delta_{2,H}$, which directly violates the definition of the high type. Second, if $\theta_L \delta_{1,L} < \theta_H \delta_{1,H}$ and since ${\pi_2}^{H,1} \geq \left(1 - {\pi_1}^{H,1}\right)$, $V_{\mathcal{M}}(E_1^H, \theta_H) \leq V_{\mathcal{M}}(E_1^H, \theta_L)$ requires that $\theta_H \delta_{1,H} - \theta_L \delta_{1,L} \geq (1 - \theta_H)\delta_{2,H} - (1 - \theta_L)\delta_{2,L}$. Then, if $\theta_H \delta_{1,H} - \theta_L \delta_{1,L} > 0 \geq (1 - \theta_H)\delta_{2,H} - (1 - \theta_L)\delta_{2,L}$, the definition of the high type is again directly violated. Lastly, if $(1 - \theta_H)\delta_{2,H} > (1 - \theta_L)\delta_{2,L}$,

$$V_{\mathcal{M}}(E_1^L, \theta_H) \leq V_{\mathcal{M}}(E_1^L, \theta_L)$$

implies that ${\pi_1}^{L,1} \geq {\pi_1}^{H,1}$ and ${\pi_2}^{L,1} \leq {\pi_2}^{H,1}$. But, this contradicts that $V_{\mathcal{M}}(E_1^H, \theta_H) < V_{\mathcal{M}}(E_1^H, \theta_L)$. The proof for when the choose different actions is analogous.

In turn, this implies that the incentive-compatibility constraint of the high type binds. Hence, the data-seller's expected profits from data-buyer 1 are

$$\Pi(\mathcal{M}_1) = V_{\mathcal{M}}(E_1^L, \theta_L) + (\nu + \rho)(V_{\mathcal{M}}(E_1^H, \theta_H) - V_{\mathcal{M}}(E_1^L, \theta_H))$$

Moreover, it is trivial that selecting the most informative E_1^L such that $V_{\mathcal{M}}(E_1^L, \theta_H)$ equals $V_{\mathcal{M}}(E_1^L, \theta_L)$ maximizes the data-seller's profits from data-buyer 1. When $c < 1$, this also increases data-buyer 2 willingness to pay for information and, therefore, weakly increases the data-seller's profits overall. When $c > 1$, using an analogous argument as before, either $\mathcal{M} \in \{\underline{\mathcal{M}}^2, \underline{\mathcal{M}}^3, \underline{\mathcal{M}}^4\}$ or the gain from data-buyer 1 in profits exceeds the loss from data-buyer 2. In either case, the participation constraint of the high type binds.

The proof for the case in which data-buyers choose different actions without supplemental information is analogous.

4. Since all participation constraints bind, incentive-compatibility requires

$$IC_{L,1} : 0 \geq V_{\mathcal{M}}(E_1^H, \theta_L) - V_{\mathcal{M}}(E_1^H, \theta_H) \text{ and } IC_{H,1} : 0 \geq V_{\mathcal{M}}(E_1^L, \theta_H) - V_{\mathcal{M}}(E_1^L, \theta_L).$$

We've shown that in any optimal menu $V_{\mathcal{M}}(E_1^H, \theta_L) < V_{\mathcal{M}}(E_1^H, \theta_H)$, implying that $IC_{L,1}$ is slacked. We also showed that it is optimal for the data-seller to select E_1^L such that $V_{\mathcal{M}}(E_1^L, \theta_H) = V_{\mathcal{M}}(E_1^L, \theta_L)$. ■

Proof. Proposition 2. Consider first the case in which data-buyers would choose the same action without additional information. Assume without loss of generality that they choose action a_1. Proposition 1 implies that the optimal personalized menu is characterized by $(\pi_2^{L,i}, \pi_2^{H,i}) \in [0,1]^2$ for $i \in \{1,2\}$, which maximize the data-seller's expected profits subject to

$$V_{\mathcal{M}}\left(E_1^L, \theta_L\right) = V_{\mathcal{M}}\left(E_1^L, \theta_H\right) \text{ and } V_{\mathcal{M}}\left(E_2^L, \theta_L\right) = V_{\mathcal{M}}\left(E_2^L, \theta_H\right),$$

where

$$V_{\mathcal{M}}\left(E_i^L, \theta_L\right) = (1-\theta_L)\,\pi_2^{L,i}\left[\frac{(1-2\nu-\rho)}{1-\nu-\rho}\left[(1-\pi_2^{L,j})c + \pi_2^{L,j}\right] + \frac{\nu}{1-\nu-\rho}\left[(1-\pi_2^{H,j})c + \pi_2^{H,j}\right]\right] \text{ and}$$
$$V_{\mathcal{M}}\left(E_i^L, \theta_H\right) = (1-\theta_H)\,\pi_2^{L,i}\left[\frac{\nu}{\nu+\rho}\left[(1-\pi_2^{L,j})c + \pi_2^{L,j}\right] + \frac{\rho}{\nu+\rho}\left[(1-\pi_2^{H,j})c + \pi_2^{H,j}\right]\right],$$

for $i, j \in \{1,2\}$ and $i \neq j$. Note that both constraints hold if and only if

$$\pi_2^{L,1} = \pi_2^{L,2} = 0 \text{ or } \pi_2^{L,1} = p_1(\pi_2^{H,1}) \text{ and } \pi_2^{L,2} = p_2(\pi_2^{H,2}),$$

where $p_1(\pi_2^{H,1}) = p_2(\pi_2^{H,2})$ if and only if $\pi_2^{H,1} = \pi_2^{H,2}$. When $\pi_2^{L,1} = \pi_2^{L,2} = 0$, the data-seller's expected profits are

$$\Pi(\mathcal{M}) = (1-\theta_H)\left[\pi_2^{H,2}(\nu \cdot c + \rho(c(1-\pi_2^{H,1}) + \pi_2^{H,1})) + \pi_2^{H,1}(\nu \cdot c + \rho(c(1-\pi_2^{H,2}) + \pi_2^{H,2}))\right].$$

Hence, it is optimal for the data-seller to set $\pi_2^{H,1} = \pi_2^{H,2} = 1$ for all $c < \frac{2\rho}{\rho-\nu}$, implying that when it is optimal for the data-seller to offer no information to both low type data-buyers, then it is optimal to offer full information to both high type data-buyers. When $c \geq \frac{2\rho}{\rho-\nu}$, the data-seller finds it optimal to offer a personalized menu characterized by $\pi_2^{H,1} = 1$ and $\pi_2^{H,2} = 0$ or a symmetric menu with $\pi_2^{H,1} = \pi_2^{H,2} = 1$.

When $\pi_2^{L,1} = p_1(\pi_2^{H,1})$, $\pi_2^{L,2} = p_2(\pi_2^{H,2})$ and $c < 2$, it is optimal for the data-seller to set $\pi_2^{H,1} = \pi_2^{H,2} = 1$, implying that $\pi_2^{L,1} = \pi_2^{L,2} \in (0,1)$. Hence, in this case, it is also optimal for the data-seller to offer the same menu to both data-buyers. Lastly, when $c \geq 2$, $\pi_2^{H,1} = 1$ and $\pi_2^{H,2} \in \{0,1\}$. Hence, whenever $\pi_2^{H,2} = 1$, the menu is symmetric and asymmetric otherwise. The optimal asymmetric menu in this case is to offer no information to one data-buyer and to offer a menu with full and no information to the other one.

Therefore, a necessary condition for the optimal menu to be personalized is $c \geq \max\{2, \frac{2\rho}{\rho-\nu}\}$. If an optimal menu is personalized, it offers no information to one data-buyer and a menu to the other data-buyer with full information for the high type and no information for the low type. Otherwise, it is optimal for the data-seller to offer the same menu to both data-buyers.

Consider now the case in which data-buyers would choose different actions without additional information. The optimal menu, characterized by $\pi_1^{L,1} \in [0,1]$, $\pi_1^{L,2} \in [0,1]$ and $\pi_2^{H,2} \in [0,1]$, maximizes the data-seller's expected profits

$$\begin{aligned}\Pi(\mathcal{M}) = \theta_L &\left[\pi_1^{L,1}((1-2\nu-\rho)[(1-\pi_1^{L,2})c + \pi_1^{L,2}] + \nu) + \pi_1^{L,2}((1-2\nu-\rho)[(1-\pi_1^{L,1})c + \pi_1^{L,1}] + \nu)\right] \\ &+ (1-\theta_H)\left[\pi_2^{H,1}(\nu + \rho[(1-\pi_2^{H,2})c + \pi_2^{H,2}]) + \pi_2^{H,2}(\nu + \rho[(1-\pi_2^{H,1})c + \pi_2^{H,1}])\right]\end{aligned}$$

subject to

$$V_{\mathcal{M}}\left(E_1^L, \theta_L\right) = V_{\mathcal{M}}\left(E_1^L, \theta_H\right) \text{ and } V_{\mathcal{M}}\left(E_2^L, \theta_L\right) = V_{\mathcal{M}}\left(E_2^L, \theta_H\right),$$

where

$$\begin{aligned}V_{\mathcal{M}}\left(E_i^L, \theta_L\right) &= \theta_L \pi_1^{L,i}\left[\frac{(1-2\nu-\rho)}{1-\nu-\rho}\left[(1-\pi_1^{L,j})c + \pi_1^{L,j}\right] + \frac{\nu}{1-\nu-\rho}\right] \\ V_{\mathcal{M}}\left(E_i^L, \theta_H\right) &= \max\left\{0, (1-\theta_H)\left[\frac{\nu}{\nu+\rho} + \frac{\rho}{\nu+\rho}\left[(1-\pi_2^{H,j})c + \pi_2^{H,j}\right]\right]\right. \\ &\left. \quad - \theta_H(1-\pi_1^{L,i})\left[\frac{\nu}{\nu+\rho}\left[(1-\pi_1^{L,j})c + \pi_1^{L,j}\right] + \frac{\rho}{\nu+\rho}\right]\right\}\end{aligned}$$

for $i \in \{1,2\}$, $i \neq j$.

When $c \leq \frac{2\rho+\nu}{\rho}$, it is optimal for the data-seller's to set $\pi_2^{H,1} = \pi_2^{H,2} = 1$ since profits are strictly increasing in both variables. This, in turn, implies that $\pi_1^{L,1} = \pi_1^{L,2}$. Otherwise, when $c > \frac{2\rho+\nu}{\rho}$, the data-seller's expected profits are maximized by setting either $\pi_2^{H,1}$ or $\pi_2^{H,2}$ to 1. Assume without loss of generality that full information is offered to data-buyer 1. In turn, this means that the data-seller's expected profits are maximized by setting $\pi_2^{H,2} = 0$. Thus, the optimal menu is symmetric if and only if $c \leq \frac{2\rho+\nu}{\rho}$.

Moreover, when the optimal menu is personalized, $\pi_2^{H,1} = 1$ and $\pi_2^{H,2} = 0$, which implies that $\pi_1^{L,2} = 0$. That is, data-buyer 1 type θ_H is offered full information, whereas data-buyer 2 is offered no information. This, in turn, implies that the optimal menu coincides with the optimal menu with one data-buyer, offering partial information to the low type of data-buyer 1. ■

Proof. Proposition 3. Proposition 1 implies that $\pi_2^L = 1$ if beliefs are non-congruent and that π_1^L is such that IC_H binds. When θ_L attaches a positive value to E^L,

$$V_{\mathcal{M}}(E^L, \theta_L) = \theta_L \pi_1^L \left[\left(\frac{1-2\nu-\rho}{1-\nu-\rho} \right) [\pi_1^L + (1-\pi_1^L)c] + \left(\frac{\nu}{1-\nu-\rho} \right) \right] \text{ and}$$

$$V_{\mathcal{M}}(E^L, \theta_H) = \max \left\{ 0, (1-\theta_H) - (1-\pi_1^L)\theta_H \left(\frac{\nu}{\nu+\rho}(\pi_1^L + (1-\pi_1^L)c) + \frac{\rho}{\nu+\rho} \right) \right\}.$$

Define $f(\pi_1^L)$ as follows:

$$f(\pi_1^L) := (1-\theta_H) - (1-\pi_1^L)\theta_H \left(\frac{\nu}{\nu+\rho}(\pi_1^L + (1-\pi_1^L)c) + \frac{\rho}{\nu+\rho} \right).$$

Note that $f(\pi_1^L)$ is a continuous function of $\pi_1^L \in [0,1]$, $f(1) = (1-\theta_H) > 0$ and $f(0) < 0$. Then, the intermediate value theorem implies that there exists $\underline{\pi_1^L} \in (0,1)$ such that $f(\underline{\pi_1^L}) = 0$. Furthermore, since the graph of $f(\pi_1^L)$ is a convex parabola, $f(\pi_1^L) \geq 0$ for all $\pi_1^L \geq \underline{\pi_1^L}$ and negative otherwise. Hence, $V_{\mathcal{M}}(E^L, \theta_H) = f(\pi_1^L)$ if $\pi_1^L \geq \underline{\pi_1^L}$ and $V_{\mathcal{M}}(E^L, \theta_H) = 0$ otherwise. Define $S^N(\pi_1^L)$ as the high type surplus from acquiring experiment E^L. That is:

$$S^N(\pi_1^L) = V_{\mathcal{M}}(E^L, \theta_H) - V_{\mathcal{M}}(E^L, \theta_L).$$

Note that $S^N(\pi_1^L = \underline{\pi_1^L}) < 0$, $S^N(\pi_1^L = 1) > 0$ because $\theta_L < 1-\theta_H$ and $S^N(\pi_1^L)$ is continuous on the closed interval $[\underline{\pi_1^L}, 1]$. Then, by the Intermediate value theorem, there exists $\pi_1^L \in (\underline{\pi_1^L}, 1)$ such that $S^N(\pi_1^L) = 0$. This implies that the low type observes partial information whenever beliefs are non-congruent. ■

Proof. Lemma 7. I first show that the precision of the optimal E^L decreases as the coordination incentives increase. Since coordination incentives decrease in c, this is equivalent to showing that the optimal π_1^L increases in c. First, note that the high type surplus from acquiring E^L, $S^N(\pi_1^L)$, is an increasing function of π_1^L since

$$\frac{\partial S^N(\pi_1^L)}{\partial \pi_1^L} \geq \theta_H \left[\frac{\nu}{\nu+\rho}[\pi_1^L + (1-\pi_1^L)(2c-1)] + \frac{\rho}{\nu+\rho}\right] \\ - (1-\theta_H)\left[\frac{1-2\nu-\rho}{1-\nu-\rho}[\pi_1^L(2-c) + (1-\pi_1^L)c] + \frac{\nu}{1-\nu-\rho}\right] \geq 0.$$

The first inequality holds since $\theta_L < 1-\theta_H$ by definition of the high type whereas the second one holds for all non-congruent distribution of types, for all c such that a symmetric menu is optimal and for all distribution of private information. Second, note that the optimal π_1^L, defined as π_1^L such that $S^N(\pi_1^L) = 0$, increases in c because $S^N(\pi_1^L)$ decreases in c. Hence, the precision of the optimal E^L decreases as the coordination incentives increase.

Second, I show that the precision of the optimal E^L decreases in the correlation of private information when data-buyers have coordination incentives. Define the degree of correlation of private information as

$$\kappa := \frac{\mathbb{P}(\theta_i = \theta_H|\theta_j = \theta_H)\mathbb{P}(\theta_i = \theta_L|\theta_j = \theta_L)}{\mathbb{P}(\theta_i = \theta_L|\theta_j = \theta_H)\mathbb{P}(\theta_i = \theta_H|\theta_j = \theta_L)} = \frac{\rho(1-2\nu-\rho)}{\nu^2},$$

where data-buyers' private information is more correlated as κ increases. In particular, by definition, data-buyers' private information is positively (negatively) correlated when $\kappa > (<)1$ and conditionally independent when $\kappa = 1$.

Note that κ is a decreasing function of ν and that the effect of ν on $S^N(\pi_1^L)$ depends on the strategic incentives because

$$\frac{\partial S^N(\pi_1^L)}{\partial \nu} = (1-\pi_1^L)(c-1)\left[\frac{1-\rho}{(1-\nu-\rho)^2}\theta_L\pi_1^L - \theta_H(1-\pi_1^L)\frac{\rho}{(\rho+\nu)^2}\right].$$

In particular, $S^N(\pi_1^L)$ is decreasing function of ν when data-buyers have coordination incentives ($c < 1$). To see this, note that $S^N(\pi_1^L)$ is an decreasing function of ν if and only if

$$\theta_H(1-\pi_1^L)\frac{\rho}{(\rho+\nu)^2} - \frac{1-\rho}{(1-\nu-\rho)^2}\theta_L\pi_1^L \leq 0. \tag{6}$$

and (6) holds for all

$$\pi_1^L \geq \frac{\theta_H\rho(1-\nu-\rho)^2}{\theta_H\rho(1-\nu-\rho)^2 + \theta_L(1-\rho)(\nu+\rho)^2} \text{ and } S^N\left(\frac{\theta_H\rho(1-\nu-\rho)^2}{\theta_H\rho(1-\nu-\rho)^2 + \theta_L(1-\rho)(\nu+\rho)^2}\right) < 0, \tag{7}$$

where the optimal π_1^L satisfies (7). Hence, the optimal π_1^L increases in ν since it is defined as π_1^L such that $S^N(\pi_1^L) = 0$ where $S^N(\pi_1^L)$ is an increasing function of π_1^L. This, in turn, implies that the optimal π_1^L decreases in κ. Analogously, we can conclude that that the optimal π_1^L increases in the correlation of private information when data-buyers have anti-coordination incentives. ■

Lemma 11 *Assume $c < 1$ and that beliefs are congruent. In an optimal menu, the data-seller offers no information to θ_L if $\nu \leq \sqrt{\rho} - \rho$ and partial information otherwise.*

Proof. Lemma 11. Proposition 1 implies that in an optimal menu $E^H = \overline{E}$, $\pi_1^L = 1$ and that π_2^L is determined such that IC_H binds. Given that both participation constraints bind, IC_H simplifies to $V_{\mathcal{M}}(E^L, \theta_L) = V_{\mathcal{M}}(E^L, \theta_H)$ where both expressions are computed by assuming that j does not deviate form her equilibrium choices. The value of experiment E^L for the θ_L and θ_H are given by

$$V_{\mathcal{M}}(E^L, \theta_L) = (1-\theta_L)\pi_2^L \left[\left(\frac{1-2\nu-\rho}{1-\nu-\rho}\right)\left[(1-\pi_2^L)c + \pi_2^L\right] + \left(\frac{\nu}{1-\nu-\rho}\right)\right] \text{ and}$$
$$V_{\mathcal{M}}(E^L, \theta_H) = (1-\theta_H)\pi_2^L \left[\left(\frac{\nu}{\nu+\rho}\right)\left[(1-\pi_2^L)c + \pi_2^L\right] + \left(\frac{\rho}{\nu+\rho}\right)\right].$$

It is trivial that IC_H binds if $\pi_2^L = 0$. Suppose now that $\pi_2^L \in (0, 1]$. Assume first that $\nu \leq \sqrt{\rho} - \rho$. Then:

$$\begin{aligned} V_{\mathcal{M}}(E^L, \theta_H) - V_{\mathcal{M}}(E^L, \theta_L) &> (1-\theta_L)\,\pi_2^L \left(\frac{(1-2\nu-\rho)}{1-\nu-\rho}[\pi_2^L + (1-\pi_2^L)c] + \frac{\nu}{1-\nu-\rho}\right) \\ &\quad - (1-\theta_L)\,\pi_2^L \left(\frac{(1-2\nu-\rho)}{1-\nu-\rho}[\pi_2^L + (1-\pi_2^L)c] + \frac{\nu}{1-\nu-\rho}\right) \\ &= 0 \end{aligned}$$

where the inequality holds since $\theta_L > \theta_H$, $c < 1$ and $\nu \leq \sqrt{\rho} - \rho$. Thus, it is not possible for the data-seller to offer partial information to the low type without inducing a deviation from the high type. Assume now that $\nu > \sqrt{\rho} - \rho$. In this case, there exists $\pi_2^L \in (0, 1]$ such that $V_{\mathcal{M}}(E^L, \theta_L) = V_{\mathcal{M}}(E^L, \theta_H)$ if and only if

$$\theta_L \leq \frac{(1-\nu-\rho)[c\cdot\nu+\rho]\theta_H + [(\nu+\rho)^2 - \rho](1-c)}{[c(1-2\nu-\rho)+\nu](\nu+\rho)} := \hat{\theta} \tag{8}$$

where

$$\pi_2^L = \frac{(1-\nu-\rho)\left[\nu c+\rho\right]\theta_H + \left[(\nu+\rho)^2 - \rho\right](1-c) - \left[c(1-2\nu-\rho)+\nu\right](\nu+\rho)\theta_L}{(1-c)\left[(\nu+\rho)^2 - \rho - \nu(1-\nu-\rho)\theta_H + (\nu+\rho)(1-2\nu-\rho)\theta_L\right]}. \tag{9}$$

Thus, if the low type is sufficiently uncertain about the state, the data-seller is able to provide supplemental information to the low type without attracting the high type. Otherwise, the low type observes no supplemental information. ∎

Lemma 12 *Assume $c > 1$ and that beliefs are congruent. In an optimal menu, the data-seller offers no information to θ_L if $\nu \geq \sqrt{\rho} - \rho$ and partial information otherwise.*

Proof. Lemma 12. The proof is analogous to the proof of Lemma 11. ∎

Proof. Proposition 4. This proof is contained in Lemma 11 and Lemma 12 where $\hat{\theta}$ is defined as the left-hand side of Equation (8). ∎

Proof. Lemma 8. In an optimal menu in which the low type observes partial information, (9) defines the optimal E^L. First, if data-buyers' private information is negatively correlated and they have coordination incentives, $c < 1$ and $\nu > \sqrt{\rho} - \rho$. In this case, the sign of $\frac{\partial \pi_2^L}{\partial c}$ depends on the sign of

$$(\nu + \rho)^2 - \rho - \nu(1 - \nu - \rho)\theta_H + (1 - 2\nu - \rho)(\nu + \rho)\theta_L$$

which is positive for all $\theta_L > \theta_H$ and $\nu > \sqrt{\rho} - \rho$. Hence, the precision of E^L increases in c. As c increases, data-buyers incentives to coordinate decrease. Thus, the precision of E^L increases as the incentives to coordinate decrease. Similarly, the sign of $\frac{\partial \pi_2^L}{\partial \nu}$ depends on the sign of

$$(\nu + \rho)(\theta_H \rho(2 - \nu - \rho) - \theta_L(1 - \rho)(\nu + \rho) + \nu - \rho) + (1 - \theta_H)\rho$$

which is positive for all $\theta_L > \theta_H$, $\nu > \sqrt{\rho} - \rho$ and $c < 1$. Hence, the precision of E^L increases in ν. An increase in ν decreases κ which measures the degree of correlation between their private information, defined in Lemma 7. Thus, the precision of E^L decreases in the correlation of private information.

Second, if data-buyers' private information is positively correlation and they have anti-coordination incentives, $c > 1$ and $\nu < \sqrt{\rho} - \rho$. Analogously, it is straightforward to show that $\frac{\partial \pi_2^L}{\partial c} > 0$ and $\frac{\partial \pi_2^L}{\partial \nu} < 0$. ∎

A.2 Proofs with N data-buyers

Lemma 13 *The value of experiment E^n is increasing in its precision if*

$$c < 1 \text{ and } c\left(1 + \left\lceil \frac{N}{2} \right\rceil\right) \geq \left\lceil \frac{N}{2} \right\rceil \text{ or } c > 1 \text{ and } \left(1 + \left\lceil \frac{N}{2} \right\rceil\right) \geq c\left\lceil \frac{N}{2} \right\rceil.$$

Proof. Lemma 13. Consider the case in which data-buyer i acquires experiment E^L. $V_{\mathcal{M}}(E^L, \theta)$ is increasing in π_1^L if

$$\Lambda_1^\theta \geq \max\left\{ -\pi_1^L \cdot \frac{\partial \Lambda_1^\theta}{\partial \pi_1^L}, (1-\pi_1^L)\frac{\partial \Lambda_1^\theta}{\partial \pi_1^L} \right\}. \tag{10}$$

A sufficient but not necessary condition for (10) is

$$\Lambda_1^\theta \geq \max\left\{ -\frac{\partial \Lambda_1^\theta}{\partial \pi_1^L}, \frac{\partial \Lambda_1^\theta}{\partial \pi_1^L} \right\} = \mid \frac{\partial \Lambda_1^\theta}{\partial \pi_1^L} \mid \tag{11}$$

since $\pi_1^L \in [0,1]$. Λ_1^θ depends on the distribution of the Conway-Maxwell-Binomial random variable, κ_{-i}^1. The distribution function of a Conway-Maxwell-Binomial (n, p, ν) random variable is given by

$$F(k; n, p, \nu) = \frac{\sum_{\ell=0}^{k} p^\ell (1-p)^{n-\ell} \binom{n}{\ell}^\nu}{S(p,\nu)}.$$

where $S(p,\nu) = \sum_{k=0}^{n} p^k (1-p)^{n-k} \binom{n}{k}^\nu$ is a normalizing constant. Given that π_1^L only affects $p_{\omega_1,\theta}$, using the chain rule, we have:

$$\frac{\partial F(k; n, p_{\omega_1,\theta}, \nu_{\omega_1,\theta})}{\partial \pi_1^L} = \frac{\partial F(k; n, p_{\omega_1,\theta}, \nu_{\omega_1,\theta})}{\partial p_{\omega_1,\theta}} \frac{\partial p_{\omega_1,\theta}}{\partial \pi_1^L}$$

where $\frac{\partial p_{\omega_1,\theta}}{\partial \pi_1^L} = \mathbb{P}(\theta_j = \theta_L | \theta)$ and

$$\frac{\partial F(k;n,p,\nu)}{\partial p_{\omega_1,\theta}} = \frac{\sum_{\ell=1}^{k} \binom{n}{\ell}^\nu p^{\ell-1}(1-p)^{n-\ell-1}(\ell - np) - \left[\sum_{\ell=1}^{n} \binom{n}{\ell}^\nu p^{\ell-1}(1-p)^{n-\ell-1}(\ell-np)\right] F(k;n,p,\nu)}{S(p,\nu)}.$$

Note that

$$\begin{aligned}\frac{\partial F(k;n,p,\nu)}{\partial p_{\omega_1,\theta}} &\leq (k-np)\frac{\sum_{\ell=1}^{k} \binom{n}{\ell}^\nu p^{\ell-1}(1-p)^{n-\ell-1}}{S(p,\nu)} + np\frac{\sum_{\ell=1}^{n} \binom{n}{\ell}^\nu p^{\ell-1}(1-p)^{n-\ell-1}}{S(p,\nu)} \\ &\leq (k-np) + np = k.\end{aligned}$$

Then, $\frac{\partial \Lambda_1^\theta}{\partial \pi_1^L}$ is given by:

$$\left|\frac{\partial \Lambda_1^\theta}{\partial \pi_1^L}\right| = \mathbb{P}(\theta_j = \theta_L|\theta)\left|(1-c)\frac{\partial F(\lceil \frac{N}{2}\rceil - 1; \cdot)}{\partial p_{\omega_1,\theta}}\right| \leq \left|(1-c)\lceil \frac{N}{2} \rceil\right|$$

where the inequality holds since $\mathbb{P}(\theta_j = \theta_L|\theta) \in [0,1]$. Moreover, $\Lambda_1^\theta \geq \min\{1, c\}$. Then, (11) holds for all $\pi_1^L \in [0,1]$ (or $p_{\omega_1,\theta}$) if

$$c \geq (1-c)\lceil \frac{N}{2} \rceil \text{ if } c < 1 \Leftrightarrow c\left(1 + \left\lceil \frac{N}{2} \right\rceil\right) \geq \left\lceil \frac{N}{2} \right\rceil$$
$$1 \geq (c-1)\lceil \frac{N}{2} \rceil \text{ if } c > 1 \;\Leftrightarrow \left(1 + \left\lceil \frac{N}{2} \right\rceil\right) \geq c\left\lceil \frac{N}{2} \right\rceil.$$

Analogously, I can show that $V_{\mathcal{M}}(E^H, \theta)$ increases in π_1^H under the same sufficient conditions and that $V_{\mathcal{M}}(E^n, \theta)$ increases in π_2^n. ■

Proposition 9 *Assume that payoffs satisfy (3). The high type observes* $\overline{E}$ *and*

1. *If* $\nu \leq 1$*, the low type observes no information if beliefs are congruent and partial information if beliefs are non-congruent.*
2. *If* $\nu > 1$*, there exists* $\tilde{\theta} \in (0,1)$ *such that the low type observes partial information if beliefs are congruent and* $\theta_L < \tilde{\theta}$*. and no information otherwise.*

Proof. Proposition 9. When beliefs are congruent, $\pi_1^L = 1$ and π_2^L such that

$$V_{\mathcal{M}}(E^L, \theta_L) = V_{\mathcal{M}}(E^L, \theta_H).$$

First, the value that the low type attaches to experiment E^L is given by:

$$V_{\mathcal{M}}(E^L, \theta_L) = (1-\theta_L)\pi_2^L \Lambda_2^{\theta_L}$$

where $\kappa_{-i}^1|(\omega_2, \theta_L)$ is distributed according to a CMB distribution with parameters $N-1$, $p_{\omega_2,\theta_L} = \eta_L(1-\pi_2^L)$ and ν. Second, the value that the high type attaches to experiment E^L is given by:

$$V_{\mathcal{M}}(E^L, \theta_H) = (1-\theta_H)\pi_2^L \Lambda_2^{\theta_H}$$

where $\kappa_{-i}^1|(\omega_2, \theta_H)$ is distributed according to a CMB distribution with parameters $N-1$, $p_{\omega_2,\theta_H} = (1-\eta_H)(1-\pi_2^L)$ and ν. Denote by $F(k; N-1, p_{\omega,\theta_i}, \nu)$ the distribution function of a CMB distribution with these parameters. Given that the distribution of $\kappa_{-i}^1|(\omega_2, \theta_L)$ and $\kappa_{-i}^1|(\omega_2, \theta_H)$ share two of those parameters, I simplify the notation to $F(k; p_{\omega,\theta_i})$.

It is trivial the incentive compatibility constraint of the high type binds if $\pi_2^L = 0$. Assume now that $\pi_2^L \in (0,1]$ and consider first the case in which data-buyers' private information is positively correlated or $\nu < 1$. In this case, $\eta_L > 1 - \eta_H$ which implies that $p_{\omega_2,\theta_L} > p_{\omega_2,\theta_H}$ and

$$V_{\mathcal{M}}(E^L, \theta_L) < (1-\theta_H)\pi_2^L \left[c + (1-c)F\left(N - \lceil \frac{N}{2} \rceil - 1; (1-\eta_H)(1-\pi_2^L)\right)\right] = V_{\mathcal{M}}(E^1, \theta_H)$$

where the inequality holds since $\theta_H < \theta_L$, $\eta_L > 1-\eta_H$ and $c < 1$. Then, the value of experiment E^L for the low type is lower than the one for the high type, $V_{\mathcal{M}}(E^L, \theta_L) < V_{\mathcal{M}}(E^L, \theta_H)$, which implies that the incentive-compatibility constraint of the high type is violated. Thus, it is not possible for the seller to offer information to the low type. Analogously, we reach the same conclusion for $\nu = 1$.

In contrast, if data-buyer types are negatively correlated ($\nu > 1$) or $\eta_L < 1-\eta_H$, we have that $(1-\theta_L) < (1-\theta_H)$ but

$$c + (1-c)F\left(N - \lceil \frac{N}{2} \rceil - 1; \eta_L(1-\pi_2^L)\right) > c + (1-c)F\left(N - \lceil \frac{N}{2} \rceil - 1; (1-\eta_H)(1-\pi_2^L)\right).$$

Define

$$\Delta V_{\mathcal{M}}(\pi_2^L) := \frac{V_{\mathcal{M}}(E^L, \theta_L) - V_{\mathcal{M}}(E^L, \theta_H)}{\pi_2^1}.$$

Note that $\Delta V_{\mathcal{M}}(1) < 0$ by the definition of high type and

$$\lim_{\pi_2^L \to 0} \Delta V_{\mathcal{M}}(\pi_2^L) \geq 0 \text{ iff } \theta_L \leq 1 - \frac{(1-\theta_H)\left(c + (1-c)F\left(N - \lceil \frac{N}{2} \rceil - 1; (1-\eta_H)\right)\right)}{c + (1-c)F\left(N - \lceil \frac{N}{2} \rceil - 1; \eta_L\right)} := \tilde{\theta}. \quad (12)$$

Then, the Intermediate value theorem implies that there exists at least one $\pi_2^L \in (0,1)$ such that $V_{\mathcal{M}}(E^L, \theta_L) = V_{\mathcal{M}}(E^L, \theta_H)$. Thus, the seller is able to offer partial information to the low type if and only if (12) holds.

Consider now the case in which beliefs are non-congruent. In this case, $\pi_2^L = 1$ and π_1^L is such that $V_{\mathcal{M}}(E^L, \theta_L) = V_{\mathcal{M}}(E^L, \theta_H)$. First, the value that the low type attaches to experiment E^L is given by $V_{\mathcal{M}}(E^L, \theta_L) = \theta_L \pi_1^L \Lambda_1^{\theta_L}$, where $\kappa_{-i}^1|(\omega_1, \theta_L)$ is distributed according to a CMB distribution with parameters $N-1$, $p_{\omega_1,\theta_L} = \eta_L \pi_1^L + (1-\eta_L)$ and ν. Second, the high type attaches a value to experiment E^L given by

$$V_{\mathcal{M}}(E^L, \theta_H) = \max\{0, (1-\theta_H)\Lambda_2^{\theta_H} - \theta_H(1-\pi_1^L)\Lambda_1^{\theta_H}\}$$

where $\kappa_{-i}^1|(\omega_1, \theta_H)$ is distributed according to a CMB distribution with parameters $N-1$, $p_{\omega_1,\theta_L} = \eta_H + (1-\eta_H)\pi_1^L$ and ν and $\kappa_{-i}^1|(\omega_2, \theta_H)$ is distributed according to a CMB distribution with parameters $N-1$, $p_{\omega_2,\theta_H} = 0$ and ν.[22] Note that $V_{\mathcal{M}}(E^L, \theta_H)$ is a continuous and increasing function of π_1^L, $V_{\mathcal{M}}(E^L, \theta_H) < 0$ if $\pi_1^L = 0$ and $V_{\mathcal{M}}(E^L, \theta_H) > 0$. Then, the intermediate value theorem implies that there exists $\hat{\pi}_1^L \in (0,1)$ such that $V_{\mathcal{M}}(E^L, \theta_H) \geq 0$ for all $\pi_1^L \geq \hat{\pi}_1^L$. Define $\Delta V_{\mathcal{M}}(\pi_1^L) = V_{\mathcal{M}}(E^L, \theta_L) - V_{\mathcal{M}}(E^L, \theta_H)$. Note that $\Delta V_{\mathcal{M}}(\pi_1^L)$ is

[22] Then, $\mathbb{P}(\kappa_{-i}^1 \leq k|\omega_2) = 1$ for all $k \geq 0$.

also a continuous function of π_1^L, $\Delta V_{\mathcal{M}}(\hat{\pi}_1^L) \geq 0$ and $\Delta V_{\mathcal{M}}(1) = (\theta_L - (1-\theta_H)) < 0$ by the definition of high type. Thus, the intermediate value theorem implies that there exists $\pi_1^L \in (\hat{\pi}_1^L, 1)$ such that $\Delta V_{\mathcal{M}}(\pi_1^L) = 0$. ■

Proposition 10 *Assume that payoffs satisfy (4). The high type observes* $\overline{E}$ *and*

1. *If* $\nu \geq 1$, *the low type observes no information if beliefs are congruent and partial information if beliefs are non-congruent.*
2. *If* $\nu < 1$, *there exists* $\tilde{\theta} \in (0,1)$ *such that the low type observes partial information if beliefs are congruent and* $\theta_L \leq \tilde{\theta}$. *and no information otherwise.*

Proof. Proposition 10. The proof is analogous to the proof of Proposition 9. ■

Proof. Proposition 5. This proof is contained in Proposition 9 and Proposition 10. ■

Proof. Proposition 6. This proof is contained in Proposition 9 and Proposition 10. ■

A.3 Proofs with correlated signals.

Proof. Proposition 7. Consider first the case in which data-buyers have congruent beliefs. In this case, data-buyers' value of experiment E^m with $m \in \{L, H\}$ is

$$V_{\mathcal{M},\psi}\left(E^m, \theta_H\right) = V_{\mathcal{M}}\left(E^m, \theta_H\right) + (1-c)\left[\theta_H \frac{\nu\psi_1^{m,L} + \rho\psi_1^{m,H}}{\nu+\rho} + (1-\theta_H)\frac{\nu\psi_2^{m,L} + \rho\psi_2^{m,H}}{\nu+\rho}\right]$$

and

$$\begin{aligned} V_{\mathcal{M},\psi}\left(E^m, \theta_L\right) &= V_{\mathcal{M}}\left(E^m, \theta_L\right) \\ &+ (1-c)\left[\theta_L \frac{(1-2\nu-\rho)\psi_1^{m,L} + \nu\psi_1^{m,H}}{1-\nu-\rho} + (1-\theta_L)\frac{(1-2\nu-\rho)\psi_2^{m,L} + \nu\psi_2^{m,H}}{1-\nu-\rho}\right]. \end{aligned}$$

In an optimal symmetric menu, it is possible to show that the high type is offered full information by following an analogous argument from Proposition 1. This, in turn, implies that $\psi_k^{H,n} = 0$ and $\psi_k^{L,H} = 0$ for all $k \in \{1, 2\}$ and $n \in \{L, H\}$. Also, analogously to Proposition 1, it is also possible to show that E^L must be concentrated, i.e., $\pi_1^L = 1$, which also requires that $\psi_1^{L,L} = 0$, and that both participation constraints and the incentive-compatibility constraint bind. Hence, the data-seller's problem is reduced to

$$\max_{(\pi_2^L, \psi_2^{L,L})} V_{\mathcal{M},\psi}\left(E^L, \theta_L\right) \text{ subject to } V_{\mathcal{M},\psi}\left(E^L, \theta_L\right) = V_{\mathcal{M},\psi}\left(E^L, \theta_H\right)$$
$$\pi_2^L \in [0,1] \text{ and } \psi_2^{L,L} \in [\underline{\psi}_2^{L,L}, \overline{\psi}_2^{L,L}]$$

where

$$\underline{\psi}_2^{L,L} := \max\{-(\pi_2^L)^2, \left(1-\pi_2^L\right)\pi_2^L - 1, -\left(1-\pi_2^L\right)^2\} \text{ and}$$
$$\overline{\psi}_2^{L,L} := \min\{1-(\pi_2^L)^2, \pi_2^L\left(1-\pi_2^L\right), 1-\left(1-\pi_2^L\right)^2\} = \pi_2^L\left(1-\pi_2^L\right).$$

Note that for any $\pi_2^L \in [0,1]$ and $\psi_2^{L,L} \in [\underline{\psi}_2^{L,L}, \overline{\psi}_2^{L,L}]$, $V_{\mathcal{M},\psi}\left(E^L, \theta_L\right)$ is strictly increasing in $\psi_2^{L,L}$ if and only if $c < 1$. Hence, when data-buyers have coordination incentives, it will be optimal for the data-seller to offer the low type an experiment that exhibits the maximal feasible correlation $\psi_2^{L,L} = \overline{\psi}_2^{L,L}$. Similarly, when data-buyers have anti-coordination incentives, it will be optimal for the data-seller to offer the low type an experiment that exhibits the minimal feasible correlation $\psi_2^{L,L} = \underline{\psi}_2^{L,L}$.

When data-buyers have coordination incentives, the data-seller offers no information to the low types if their private information is positively correlated since this is the only experiment that satisfies incentive-compatibility for all feasible $\psi_2^{L,L}$ and π_2^L. In contrast, when their private information is negatively correlated, the data-seller offers a partially informative experiment to the low type where π_2^L is pinned down by the incentive-compatibility constraint and $\psi_2^{L,L} = \overline{\psi}_2^{L,L}$.

When data-buyers have anti-coordination incentives, the data-seller offers partial information to the low type if and only if their private information is positively correlated. In contrast, the data-seller cannot offer partial information without inducing a deviation from the high type when data-buyers private information is negatively correlated.

The proof of the non-congruent case follows the same steps and shows that it is optimal for the data-seller to set $\psi_2^{L,L} = \overline{\psi}_2^{L,L}$ when $c < 1$ and $\psi_2^{L,L} = \underline{\psi}_2^{L,L}$ when $c > 1$ and that the optimal $\pi_2^L \in (0,1)$. This implies that the low type is offered partial information and the correlation between experiments is maximized with coordination incentives and minimize with anti-coordination incentives. ■

A.4 Proofs with continuum of types

Proof. Lemma 9. I show this by contradiction. First, assume the optimal menu $\mathcal{M}^*$ is such that there exist a set $\Theta^- \subset [\theta_i^*, 1]$ of non-zero measure such that $\pi_1^*(\theta_i) < 1$ for all $\theta_i \in \Theta^-$ and $\pi_1^*(\theta_i) = 1$ for all $\theta_i \in [\theta_i^*, 1] \setminus \Theta^-$. Consider an alternative menu $\mathcal{M}'$ that replaces $\pi_1^*(\theta_i)$ with 1 for all $\theta_i \in \Theta^-$. Define

$$t^{\mathcal{M}'}(\theta_i) = t^{\mathcal{M}^*}(\theta_i) + V_{\mathcal{M}'}\left(q(\theta_i), \theta_i\right) - V_{\mathcal{M}^*}\left(q(\theta_i), \theta_i\right)$$

and note that the participation and incentive-compatibility constraints for $\mathcal{M}'$ are implied by the optimality of $\mathcal{M}^*$. The change in the data-seller's profits is

$$\begin{aligned}\int_0^1 \left(t^{\mathcal{M}'}(\theta_i) - t^{\mathcal{M}^*}(\theta_i)\right) \mathrm{d}F(\theta_i) &= \int_0^1 \left(V_{\mathcal{M}'}\left(q(\theta_i), \theta_i\right) - V_{\mathcal{M}^*}\left(q(\theta_i), \theta_i\right)\right) \mathrm{d}F(\theta_i) \\ &= (1-c) \int_0^{\theta_i^*} \theta_i \pi_1^*(\theta_i) \int_{\theta_j \in \Theta^-} f_{\theta_i}(\theta_j) \left(1 - \pi_1^*(\theta_j)\right) \mathrm{d}\theta_j \mathrm{d}F(\theta_i) \\ &+ \int_{\theta_i \in \Theta^-} \theta_i (1 - \pi_1^*(\theta_i)) \int_0^1 f_{\theta_i}(\theta_j) \left(c\left(1 - \pi_1^*(\theta_j)\right) + \pi_1^*(\theta_j)\right) \mathrm{d}\theta_j \mathrm{d}F(\theta_i)\end{aligned}$$

Therefore, the seller's profits are higher with menu $\mathcal{M}'$ for all $c < 1$, contradicting the optimality of $\mathcal{M}^*$. Moreover, there exists $\hat{c} \geq 1$ such that the data-seller prefers $\mathcal{M}'$. The proof for $\pi_2(\theta_i) = 1$ for all $\theta_i < \theta_i^*$ follows analogous steps.

Now we show the monotonicity of experiments. Consider first $\pi_2^*(\theta_i)$. From Lemma 9 we know that $\pi_2^*(\theta_i) = 1$ for all $\theta_i < \theta_i^*$ and $\pi_1^*(\theta_i) = 1$ for all $\theta_i \geq \theta_i^*$. For $\theta_i > \theta_i^*$, I show the result by contradiction. Suppose in an the optimal menu, $\pi_2^*(\theta_i)$ is increasing and consider θ_i' and θ_i'' such that $\theta_i' > \theta_i'' > \theta_i^*$. Note that

$$\begin{aligned}V_{\mathcal{M}}(q(\theta_i'), \theta_i') &= (1-\theta_i')\pi_2^*(\theta_i') \int_0^1 f_{\theta_i}(\theta_j)(c\left(1 - \pi_2^*(\theta_j)\right) + \pi_2^*(\theta_j))\mathrm{d}\theta_j \\ &< (1-\theta_i')\pi_2^*(\theta_i'') \int_0^1 f_{\theta_i}(\theta_j)(c\left(1 - \pi_2^*(\theta_j)\right) + \pi_2^*(\theta_j))\mathrm{d}\theta_j \\ &= V_{\mathcal{M}}(q(\theta_i''), \theta_i')\end{aligned}$$

Similarly, $V_{\mathcal{M}}(q(\theta_i''), \theta_i'') < V_{\mathcal{M}}(E(\theta_i'), \theta_i'')$. However, this contradicts the participation and incentive-compatibility constraints. Hence, $\pi_2^*(\theta_i)$ must be non-increasing. The result for $\pi_1^*(\theta_i)$ follows analogously. ∎

Proof. Lemma 10. The proof is analogous to the proof of Lemma 2 from Bergemann et al. (2018). ∎

Example 1 *To show that a menu of experiments with at most two informative signals is not longer optimal for the data-seller, assume that F is uniform and consider the set of information structures that are piece-wise linear, defined as follows:*

$$\pi_1(\theta_i) = \begin{cases} 0 & if\ \theta_i < \theta_0 \\ b_1\theta + d_1 & if\ \theta_i \in [\theta_0, \theta_1) \\ 1 & if\ \theta_i \geq \theta_1 \end{cases} \quad and\ \pi_2(\theta_i) = \begin{cases} 1 & if\ \theta_i < \theta_2 \\ -b_2\theta + d_2 & if\ \theta_i \in [\theta_2, \theta_3) \\ 0 & if\ \theta_i \geq \theta_3 \end{cases}$$

where $b_1 \in \left[0, \frac{1}{\theta_1}\right]$, $b_2 \in \left[0, \frac{1}{1-\theta_2}\right]$, $d_1 \in [0,1]$, $d_2 \in [0, \frac{1}{1-\theta_2}]$ *and* $0 \leq \theta_0 \leq \theta_1 \leq \theta^* \leq \theta_2 \leq \theta_3 \leq 1$. *The value of experiment* $E(\theta_i')$ *for data-buyer* i *of type* θ_i *is given by*

$$V_{\mathcal{M}}\left(E(\theta_i'), \theta_i\right) = \theta_i \pi_1(\theta_i') \left[c(\theta_1 - \theta_0) + (1-c)\left[d_1(\theta_1 - \theta_0) + \frac{b_1}{2}(\theta_1^2 - \theta_0^2) + 1 - \theta_1\right]\right]$$
$$- (1-\theta_i)(1 - \pi_2(\theta_i'))\left[\theta_2 + c(\theta_3 - \theta_2) + (1-c)\left[d_2(\theta_3 - \theta_2) - \frac{b_2}{2}(\theta_3^2 - \theta_2^2)\right]\right]$$

for all $\theta_i < \theta^*$ *and*

$$V_{\mathcal{M}}\left(E(\theta_i'), \theta_i\right) = (1-\theta_i)\pi_2(\theta_i')\left[\theta_2 + c(\theta_3 - \theta_2) + (1-c)\left[d_2(\theta_3 - \theta_2) - \frac{b_2}{2}(\theta_3^2 - \theta_2^2)\right]\right]$$
$$- \theta_i(1 - \pi_1(\theta_i'))\left[c(\theta_1 - \theta_0) + (1-c)\left[d_1(\theta_1 - \theta_0) + \frac{b_1}{2}(\theta_1^2 - \theta_0^2) + 1 - \theta_1\right]\right]$$

for all $\theta_i \geq \theta^*$. *The data-seller selects* $(b_k, d_k)_{k=1}^2$ *and* $\{\theta_k\}_{k=0}^3$ *to solve the following maximization problem:*

$$\max \int_0^1 t(\theta_i)\mathrm{d}\theta_i \text{ subject to } V_{\mathcal{M}}\left(E(\theta_i), \theta_i\right) \geq t(\theta_i) \text{ for all } \theta_i \in [0,1] \text{ and}$$
$$V_{\mathcal{M}}\left(E(\theta_i), \theta_i\right) - t(\theta_i) \geq V_{\mathcal{M}}\left(E(\theta_i'), \theta_i\right) - t(\theta_i') \text{ for all } \theta_i, \theta_i' \in [0,1]$$

First, note that no information is offered to all data-buyers of type $\theta_i \in [0, \theta_0) \cup [\theta_3, 1]$ *and full information is offered to* $\theta_i = \theta^*$. *Second, from local incentive-compatibility we can pin down the transfers* $t(\theta_i)$ *as a function of* $(b_k, d_k)_{k=1}^2$ *and* $\{\theta_k\}_{k=0}^3$.

A menu with at most two informative experiments requires that $b_1 = b_2 = 0$, $d_k = 0$ *and* $d_{k'} > 0$ *for* $k, k' \in \{1, 2\}$, $k \neq k'$. *However, as illustrated below, there is no* $d_k' \in (0,1]$ *and* $\{\theta_k\}_{k=0}^3$ *such that setting* $b_1 = b_2 = 0$ *and* $d_k = 0$ *is optimal for the data-seller. In particular, the optimal piece-wise linear information structures are listed in the following table.*

c	b_1	d_1	b_2	d_2	θ_0	θ_1	θ_2	θ_3
1/4	1.73	0	0.37	1.77	0.06	0.57	1	1
1/2	1	0.01	0.02	0.76	0	0.55	1	1
3/4	0.62	0	0.08	1.82	0	0.53	1	1
5/4	0.02	0	0.08	0.12	0	0.47	1	1

Hence, it is not sufficient for the data-seller to offer at most two informative experiments.