Submitted to *Management Science*
manuscript

Bargaining Over Data and Analytics: Sellers, Buyers and Consultants

(Authors' names blinded for peer review)

The explosive growth of online commerce has generated large quantities of data that can be used by firms to improve decision making. Some of the data collected can be directly used by firms, for example, in an advertising campaign to target certain users. In other cases, these data sets can be further analyzed using sophisticated analytic methods to substantially increase their value. These data sets are a growing source of revenue for their owners – one that can generate millions of dollars each year. We examine the exclusive selling of unique, proprietary data where the value to the data buyer can be enhanced with the use of analytic services obtained from a consultant. Because our context emphasizes the exclusive selling of proprietary and unique data (rather than general purpose data that can be sold to many buyers), announcing a fixed price for the data is not a viable option for the seller. Thus we use a Nash bargaining framework where the negotiations always involve the data seller and data buyer and could sometimes involve a consultant. Data sellers can choose to sell the data alone or sell a bundle that includes the data and complementary analytic services. Data buyers can choose to simultaneously negotiate the price of data and the price of analytic services. Alternatively the data buyer could perform these negotiations in two separate steps. Our analysis shows that the contribution of the consultant's analytic services is critical to both the seller's decision to bundle and the buyer's choice over how best to structure the negotiations. Data sellers have a natural advantage over external consultants when they choose to bundle data with complementary analytic services.

Key words: Nash Bargaining, Bundling Data with Analytic Services, Simultaneous versus Sequential Negotiations.

---

1. Introduction

The volume of data being generated continues to grow exponentially, with 180 zettabytes estimated by 2025 (Holst 2021). This is expected to increase the demand for data analytic services, and an estimated 150 zettabytes are projected to be in need of analysis by 2025 (Kulkarni 2019). Many companies have realized that the benefits of the data they collect go well beyond improvements internal to the firm. Monetizing the data – either by providing insight-driven services to other companies, or by providing access to the data outright – opens up new sources of revenue, while allowing them to potentially form relationships that are beneficial to both parties (Lotame 2020). A 2019 Forrester survey of 3,417 data and analytics decision makers reported that 54% provide an API to the data for systematic or real-time access, and 38% sell an application that enables users to see trends and insights in the data (Belissent 2020).

The process of data sharing takes multiple forms. Kroger sells purchase data captured through its loyalty card as a service to consumer packaged-goods companies such as Proctor & Gamble and Nestlé (Wixom and Ross 2017). Some companies sell data analytic solutions by applying proprietary algorithms to provide customized, actionable insights. For example, Wixom and Ross (2017) note that State Street Global Exchange combines their existing data and analytic capabilities with new research to develop information-based solutions that clients can buy. Data exchange platforms like Snowflake and Bombora also provide avenues to reach out to multiple buyers (Deichmann et al. 2016).

The specific problem considered in this paper is motivated by a data sharing problem faced by a Global Distribution System (GDS). A GDS is the base reservation system typically used by travel agents and intermediaries to enable services like the reservation of airline tickets, hotel rooms, and rental cars. These companies collect large volumes of customer booking and shopping data that is both proprietary and unique. Such data is of value to airlines and hotel chains as it allows them to make more informed decisions on flight schedules, routes, room capacities, etc.. Since each GDS tends to dominate in a specific geographical region with minimal overlap, the dataset collected by them is unique, and unavailable elsewhere. Consequently negotiation becomes an essential tool to sell their data to airlines. The negotiation framework is widely used in selling personal data in private data markets (Jung et al. 2019). Price negotiation is also a common practice when selling requires exclusive rights such as patents and property rights (Gans et al. 2008, Walden 2005). It is recognized that providing data-based insights generates more revenue than selling raw data alone (e.g., Banerjee et al. 2011, Gandhi et al. 2018). This is underscored by the fact that airlines are known to hire consultants to provide data analytic services (Henry 2017). Consequently, the GDS could potentially be better off by providing analytic services bundled with the data, and this is one of the key decisions we consider in this paper. Specifically, we consider a firm with data analytic capabilities that needs to decide between two options to sell their proprietary data – (i) sell only the data, or (ii) sell the data bundled with consultancy services – i.e., a "data product."

Buyers often hire consultants as they do not have the expertise required to analyze the data; this is underscored by Gartner, who recommend the hiring of a suitable external data and analytic service provider (Radhakrishnan et al. 2020). This practice is common – for example, Daimler Trucks Asia hired Deloitte, who "developed innovative advanced analytics techniques to proactively sense very early signals of quality and safety problems" (Deloitte 2018). In the GDS example, the booking data collected can be analyzed relatively easily by a commodity service provider however, search data is more complex and requires considerable domain expertise to analyze. Therefore, we consider two types of consultants in this paper – one who negotiates their charges for services (i.e., a specialized consultant), and another who charges a fixed price. If the seller chooses to sell just the data, the buyer will need to decide on the type of consultant to hire. And if the decision is to opt for a negotiating consultant the buyer has another decision to make – whether to include the consultant in the negotiation process with the seller (where the price for the data and the consultancy services are simultaneously determined), or to have two separate negotiations, one with the seller on the price of the data, and another with the consultant on the price for their services. The seller needs to decide on the best option incorporating the choices available to the buyer – that is, the buyer's choices (of selecting the consultant type and of subsequently choosing either a simultaneous or a sequential negotiation process if a negotiating consultant is selected) have to be embedded into the seller's decision process. As sellers of data can have multiple potential buyers, we assume that they have an alternative fallback option. Any negotiation with the buyer has to result in the seller receiving at least this outside option for the negotiation to be successful (this value could be zero if an outside option is not available).

This paper makes several contributions. We find that it is possible for the seller to be better off selling just the data despite having their own consultancy capabilities. It is also possible for the seller to be better off selling a data product (i.e., bundling the data with consultancy services) even when their services are inferior to that of the third-party consultant. This implies that if the seller is considering the development of a division to provide data analytic services to buyers, they do not need to match the capabilities of third-party consultants. Interestingly we find that the decisions of the seller and buyer are aligned with regard to the choice of data versus data product – that is, the seller's decision to sell the data as a standalone product or bundled with consultancy services is also the preferred option for the buyer. A simultaneous (three-way) negotiation helps the buyer extract more of the consultant's contribution vis-á-vis a sequential negotiation. However, leveraging the expertise of the consultant in the negotiation with the seller has a downside, as the consultant can capture a share of the value of the data. Consequently, unless the consultant's contribution is substantial relative to the value of the data, the buyer will prefer a sequential negotiation and keep the consultant from claiming any of the intrinsic value of the data. We also find that buyers could prefer to hire a fixed price consultant even when the value they add is lower than that of the negotiating consultant. Not surprisingly, a higher outside option gives more flexibility to the seller.

The price negotiation model in our paper is structurally different from other pricing models used in the extant literature such as mechanism design (Mehta et al. 2021), auction (Ghosh and Roth 2011), and query-based pricing (Koutris et al. 2015). While prior literature considers a fixed price for the data, we use a cooperative bargaining model developed by Nash (1950) where both buyer and seller mutually agree on a price. This framework provides interesting results based on the players' incentives. The paper is organized as follows. A review of relevant literature is in Section 2. Section 3 introduces the problem context and develops models associated with the negotiation framework. Sections 4 and 5 focus on the decisions of the buyer and the seller respectively. A special case of making decisions when both types of consultants contribute the same value is discussed in Section 6, and Section 7 concludes.

2. Literature Review

In this section, we present literature relevant to our work – particularly in the context of selling data as information goods, and the role of a consultant in selling information products, negotiation, and the bundling of information goods.

2.1. Selling Information Goods

Data is an experience good, and that affords data sellers a variety of pricing strategies depending on context. There is a growing body of literature on the sale of data to intermediaries who use it to tailor their products, usually for heterogeneous customers. Sundararajan (2004) identifies the optimal mix of unlimited fixed-fee and usage-based pricing for information goods. Bergemann and Bonatti (2015) study how a data provider can price queries about consumer-level information (cookies) and find that the price decreases with the size of the database and increases with the fragmentation of data sales. Bergemann et al. (2022) argue that selling additional data enables more accurate price discrimination, which however reduces all consumers' welfare. A similar result is obtained by Bimpikis et al. (2019) who show that it is optimal for the provider to sell higher precision information products at higher prices to more efficient firms. Mehta et al. (2021) show that under certain conditions, a simpler price-quantity mechanism is optimal even when the data seller allows individual buyers to filter and select the records that are of interest to them.

A related stream studies the trading of data in a two-sided data market. For example, Kushal et al. (2012) compare two simple pricing models for data and establish a pricing scheme under arbitrage and competition. Agarwal et al. (2019) propose a mathematical model to design a data marketplace while taking into account associated challenges like incentivizing buyers to reveal their true valuations, and pricing correlated datasets. Bhargava et al. (2020) develop heuristics to sell goods like sales-leads data, where buyers can either have shared or exclusive access to datasets.

Our paper differs from these both in its context and in the models developed. Our context is one where the parties involved – for example, a GDS and an airline – negotiate to arrive at prices. Babaioff et al. (2012) consider the sale of information by a monopolistic data owner to a single buyer with private information, and derive conditions under which optimal revenue is achieved with a one-round revelation mechanism. Ray et al. (2020) consider a negotiated sharing arrangement between a data owner and a buyer who has the capability to analyze the data internally, and identify conditions under which the seller should offer a demonstration to a buyer. While the B2B context studied in these papers is similar to that of our paper, our problem is quite different. We consider a buyer who needs external help to analyze the data – either from the seller or a third-party – and identify strategies for the seller (and for the buyer, given the seller's decision).

2.2. Data Consultants

Companies – even those with experience – often find it difficult to perform data analytics without professional assistance (Deal 2018). As case in point, Najjar and Kettinger (2013) discuss a major U.S. drug retailer who hired a third-party data analytics firm to host a cloud-based portal to provide analytical insights for it's suppliers. The need for consultants to assist with purchasing decisions has been recognized for a long time (e.g., Montgomery 1987, Gable 1991). For example, Ferme (1987) discusses how a consultant could assist with such a purchase. Through a series of case studies and a survey, Gable (1991) identifies six dimensions of client success when engaging an external consultant to assist with the selection of a computer-based information system. A competent consultant can identify the client's needs and act as a facilitator in the purchasing process particularly when the purchasing is done though a complicated bargaining process. We consider a setting where the buyer can hire a consultant just to analyze the data and provide insights, or to be actively involved from the data purchase negotiation stage itself.

2.3. Negotiation

We adopt the bargaining framework introduced by Nash (1950). We use both sequential and simultaneous bargaining involving the three players. Early literature on multi-lateral bargaining has considered simplistic three-player/three-cake problems in which individuals and the coalition consisting of all three players earn nothing, so that only two-player coalitions are profitable to form (Binmore 1985, Bennett et al. 1995). Rochford (1984) has shown that the allocations in a symmetric pair-wise bargaining always exist. Burguet and Caminal (2011) consider each negotiating pair with assigned beliefs about the success of their alternative negotiation with the third player.

We allow for the seller to have an outside option; the role of an outside option has been studied in a bargaining setting where it has been shown to cancel out the effect of obstinacy (Compte and Jehiel 2002, Ponsatí and Sákovics 1998, Binmore 1985). Binmore (1985) is one of the first proponents of an outside option in the bargaining context; this is developed further by Binmore et al. (1989) into an 'outside-option principle' where the outside option is only used to restrict the set of solutions. Bennett (1997) models the outside option of each player in bilateral bargaining to be the maximum utility a player can obtain by negotiating with the third player. Li et al. (2006) consider uncertainty in bilateral negotiation and show that the utility of a negotiator improves significantly when outside options are available. We consider the exogenous outside option as the lower bound of the payoff available to the seller.

2.4. Bundling Information Goods

Bundling can increase value, and one of the options the seller has is to bundle analytics and associated insights with the data – i.e., to provide a data product. Bakos and Brynjolfsson (1999) analyze optimal bundling strategies for different customer segments with various types of correlated information goods. Bakos and Brynjolfsson (2000) incorporate the "predictive value of bundling," i.e., the idea that it will be easier for a seller to predict a consumer's valuation for a collection of goods than the separate valuations of each piece in that collection. They show that pricing a bundle appropriately can allow sellers to deter entrants from selling a product that competes against one in the bundle even when the entrant has a superior cost structure or quality. This result resonates with our finding that it can be profitable for the data seller to combine consultancy services with the data (rather than to sell just the data) when the potential consultant available to the buyer can provide better analytic services than the service that comes with the seller's data product. We also find however, that under some conditions, the seller can actually leverage the consultant's contribution, and would prefer not to compete with the consultant. Geng et al. (2005) find that bundling may not be profitable when the consumer's valuations of subsequent information goods decrease quickly. We find that the buyer gets additional value from consultancy services irrespective of whether it is provided by the seller or by a third party. The decision around selling the data product vis-á-vis just the data is governed by the interplay between the added value these two consultancy options provide; the seller will sell the data alone if their consultancy contribution is substantially lower than that of the third party.

Other studies on combining information goods consider how bundling can increase consumer surplus when consumers have independent linear demands (Salinger 1995); how customized bundling is more efficient than pure bundling or individual sale (Wu et al. 2008); how bundling a buyer's requests in a data market can decrease the expected payments (Gkatzelis et al. 2015). Our unique setting – where the buyer plays an active role in decision making – sets this paper apart from others in the extant literature. In our context, the buyer not only negotiates the price with the seller and the consultant, but also decides the negotiation framework in the presence of a consultant – they can chose either to bring the consultant to the negotiating table with the seller (and have a "simultaneous" negotiation), or to have two separate negotiations, one with the seller, and another with the consultant. This makes the decision more challenging for the seller.

3. The Framework and Associated Models

In this section we introduce our framework, discuss related assumptions, and present various models associated with the framework. These models are then used to arrive at strategic decisions for the buyer and seller in later sections.

We consider a context where the owner of a unique, proprietary dataset negotiates with a potential buyer for whom the dataset has value. The negotiation, if successful, results in an exclusive contract. The value V of the dataset is not known to either party, but both know its underlying distribution. Φ represents the set of possible values, with each element $i \in \Phi$ having associated with it a value v^i which has a probability ϕ^i of being realized. We show later that it is sufficient to know an estimate of V; this is helpful as obtaining an estimate is easier than having knowledge of the underlying value distribution. We assume that the seller has available an outside option for the data of value r that both parties are aware of, and that a sale will result only if the buyer is willing to pay at least r. We assume that the negotiation is constrained by the outside option in the sense that any negotiation has to provide the seller at least the outside option irrespective of whether the negotiation is only for the data, or for the data product. We note that this is a one-time exclusive selling of the data to the buyer. We also assume that the seller has the ability to create a data product and can provide custom insights, while the buyer does not.

In the event that the seller opts to sell only the data, the buyer can choose between a specialized consultant who will negotiate to arrive at the service charges (type N), and one who charges a fixed price for their services (type F). A type N consultant can potentially be brought to the negotiating table with the seller, resulting in a simultaneous 3-way negotiation, or can be negotiated with separately. If the buyer decides not to include the consultant in the data buying process, the price of the dataset is decided through a negotiation between the seller and the buyer, with the buyer and consultant negotiating separately for the charge associated with the consultant's services, in what we refer to as a "sequential" negotiation. We note that the two negotiations do not have to happen in chronological order – i.e., the buyer can first negotiate the service charges with the consultant contingent on successfully purchasing the data, and then negotiate on the price of the data with the seller. However, irrespective of the sequence of the two negotiations, the negotiation between buyer and consultant becomes relevant only if the negotiation between buyer and seller is successful. On the other hand, the negotiation between the buyer and the seller is not dependent on the success of the negotiation between the buyer and the consultant. Therefore, the seller always has the dependency advantage over the consultant in a sequential negotiation, and that is reflected in the equilibrium outcome as discussed later.

As already mentioned, the buyer also has the option of involving the type N consultant in the price negotiation process. In the resulting simultaneous negotiation, all three players jointly decide the price for the dataset and the price of the services provided by the consultant (both of which are paid by the buyer). If the buyer decides to hire a type F consultant on the other hand, both the buyer and the seller negotiate on the price of the dataset while considering the fixed charges of the consultant.

The seller is aware that the buyer will need external help to analyze the data – that is, that one of the two types of consultants will be hired if the data is sold without analytic services. This provides the seller the option to bundle analytic services with the data, and eliminate the external consultant altogether. The seller therefore has to compare the payoffs from selling stand-alone data and the data product to make an informed decision. The buyer has to decide the type of consultant to hire (type N or type F) when only data is sold, and the negotiation format to implement (sequential or simultaneous). Figure 1 presents the decision process, while the variables and parameters used in the model are provided in Table 1.

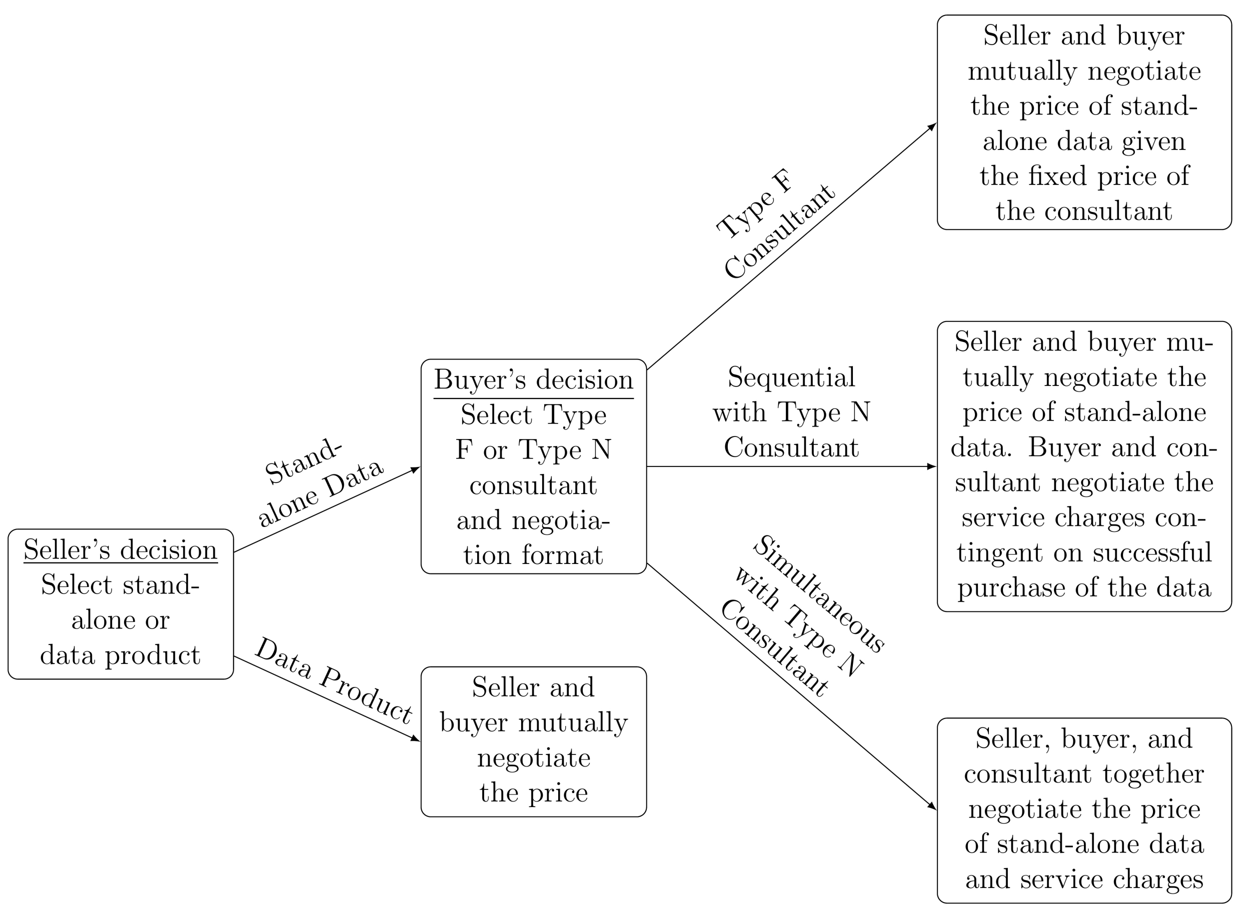


Figure 1 Decision Process

Parameters			
V	Estimated value of stand-alone data	r	Seller's outside option
δ_C	Estimated value-added contribution by type N consultant	δ_C^F	Estimated value-added contribution by type F consultant
c_C	Processing cost of type N consultant	q_C^F	Fixed price charged by type F consultant
λ_C	$= \delta_C - c_C$, Estimated net contribution by type N consultant	λ_C^F	$= \delta_C^F - q_C^F$, Estimated net contribution by type F consultant
δ_S	Estimated value-added contribution by seller	λ_S	$= \delta_S - c_S$, Estimated net contribution by seller
c_S	Processing cost of seller		
Decision Variables			
q_C	Negotiated price paid to type N	q_S	Negotiated price paid to seller

Table 1 Model Parameters and Variables

3.1. Sequential Negotiation: Type N Consultant

We start by solving the negotiation between the buyer and a type N consultant, contingent upon the successful outcome of data-purchase negotiation between the buyer and the seller. The consultant adds value by analyzing the purchased data, and we assume that all participants know the distribution of this added value. This distribution is represented by Ψ, the set of possible incremental contributions by the type N consultant, with each element $j \in \Psi$ having associated with it a contribution Δ^j which has a probability ψ^j of being realized. We assume that both $\{v^i\}_{i\in\Phi}$ and $\{\Delta^j\}_{j\in\Psi}$ are independently distributed, and that the cost c_C of processing and analyzing the data incurred by the consultant is common knowledge. If the negotiated price of data analytic services is q_C, the payoff to the buyer and the consultant for each $j \in \Psi$ are $u_B^j = \Delta^j - q_C$ and $u_C^j = q_C - c_C$ respectively. We use the bargaining framework of Nash (1950) to formulate the negotiation problem (P_{BC}) between the buyer and the consultant where the decision variable is q_C, the price paid by the buyer to the consultant for their services.

$$
\begin{aligned}
(P_{BC}) \quad & \underset{q_C}{\text{maximize}} \quad && \left(\sum_{j\in\Psi}\psi^j u_B^j\right)\cdot\left(\sum_{j\in\Psi}\psi^j u_C^j\right) && \text{(1a)}\\
& \text{subject to} && \sum_{j\in\Psi}\psi^j u_B^j \geq 0, && \text{(1b)}\\
& && \sum_{j\in\Psi}\psi^j u_C^j \geq 0 && \text{(1c)}
\end{aligned}
$$

The objective function (1a) is the Nash product obtained by multiplying the expected payoffs of the players. This is in accordance with the von Neumann-Morgenstern utilities used in the Nash product which satisfy expected utility assumptions (Rubinstein et al. 1992). Maximizing the Nash product identifies a unique solution on the Pareto-efficient frontier formed by the expected payoffs of the players. The individual rationality constraints for the buyer and the consultant are given by (1b) and (1c) respectively. In the case of a disagreement, both players will receive zero payoff. The expected payoffs of the buyer and the consultant can be reduced to $\sum_{j\in\Psi}\psi^j u_B^j = \sum_{j\in\Psi}\psi^j\Delta^j - q_C = \delta_C - q_C$ and $\sum_{j\in\Psi}\psi^j u_C^j = q_C - c_C$ respectively, where δ_C is the expected value-added contribution by the type N consultant. The negotiation problem (P_{BC}) can now be simplified as below.

$$
\begin{aligned}
(P_{BC}) \quad & \underset{q_C}{\text{maximize}} \quad && (\delta_C - q_C)\cdot(q_C - c_C) && \text{(2a)}\\
& \text{subject to} && \delta_C - q_C \geq 0, && \text{(2b)}\\
& && q_C - c_C \geq 0 && \text{(2c)}
\end{aligned}
$$

Appendix A provides the solution to this negotiation; the solution is $q_C^* = \frac{1}{2}(\delta_C + c_C)$ when $\delta_C \geq c_C$. Both the seller and the buyer will use these outcomes in the first stage of the sequential negotiation to find the price q_S for the dataset. The payoff to the buyer for each $i \in \Phi$ and $j \in \Psi$ is given by $u_B^{ij} = v^i + \Delta^j - q_S - q_C^*$. The buyer's payoff is assumed to be linear and additive in value v^i of the stand-alone data and the contribution Δ^j. This follows the linearity property of the utility function in the Nash product (Nash 1950). Assuming independence of the two distributions, we can write the buyer's expected payoff as $\sum_{i\in\Phi}\sum_{j\in\Psi}\phi^i\psi^j u_B^{ij} = \sum_{i\in\Phi}\phi^i v^i + \sum_{j\in\Psi}\psi^j\Delta^j - q_S - q_C^* = V + \delta_C - q_S - q_C^*$ where V is the expected value of the stand-alone data. The seller's payoff is the negotiated price $u_S = q_S$.

We can now express the bargaining game between the buyer and the seller in terms of the estimates V and δ_C as follows.

$$(P_{BS}) \quad \underset{q_S}{\text{maximize}} \quad q_S \cdot (V + \delta_C - q_S - q_C^*) \tag{3a}$$

$$\text{subject to} \quad q_S - r \geq 0, \tag{3b}$$

$$V + \delta_C - q_S - q_C^* \geq 0 \tag{3c}$$

We note that the seller will negotiate only if they expect to get at least r, as expressed by the individual rationality constraint (3b). The solution of the sequential negotiation is stated in Lemma 1 and derived in Appendix A.

LEMMA 1. *In a sequential negotiation, the equilibrium expected payoffs U_S^S, U_B^S, and U_C^S of the seller, buyer, and consultant respectively are as below, where $\lambda_C = \delta_C - c_C \geq 0$ is the net contribution by the consultant.*

$$(U_S^S, U_B^S, U_C^S) = \begin{cases} \left(\frac{1}{2}(V + \frac{1}{2}\lambda_C), \frac{1}{2}(V + \frac{1}{2}\lambda_C), \frac{1}{2}\lambda_C\right), & \text{if } r \leq \frac{1}{2}(V + \frac{1}{2}\lambda_C) \\ \left(r, V + \frac{1}{2}\lambda_C - r, \frac{1}{2}\lambda_C\right), & \text{if } \frac{1}{2}(V + \frac{1}{2}\lambda_C) \leq r \leq V + \frac{1}{2}\lambda_C \\ (r, 0, 0), & \text{if } V + \frac{1}{2}\lambda_C \leq r \end{cases}$$

Interestingly, both the buyer and the seller are able to extract a part of the consultant's net contribution $\left(\frac{1}{4}\lambda_C\right)$ when the outside option is small. This is because the seller knows that the buyer's negotiation with the consultant is relevant only if the data is purchased, and can take advantage of the fact that the consultant's payoff is contingent on the successful outcome of the data purchase negotiation. However, the consultant cannot claim any part of the value V as she is not involved in the data purchasing process. V has a greater share than λ_C in the equilibrium payoffs of the seller and the buyer. Therefore a large V always benefit the seller and the buyer while the consultant remains unaffected. The extra payoff of $\frac{1}{4}\lambda_C$ is particularly important for the seller as it results when the outside option r is relatively small $\left(\text{i.e., } r \leq \frac{1}{2}(V + \frac{1}{2}\lambda_C)\right)$. The negotiation fails if the outside option is large, as the buyer is not willing to pay that high an amount. Later, we will explain how the relative importance of V and λ_C in the payoff plays a crucial role for the buyer in their decision between a sequential and a simultaneous negotiation.

3.2. Simultaneous Negotiation: Type N Consultant

As already noted, firms often seek the advice of consultants in the information goods purchasing process (Montgomery 1987, Gable 1991). As the consultant is going to analyze the data after purchase, including the consultant in the price negotiation process could be helpful to the buyer. If they do, the seller, the buyer, and the consultant will jointly decide the price q_S for the dataset and the price q_C for the data analytic services. As shown in Section 3.1, the buyer's expected payoff can be expressed in terms of the expectations V and δ_C. The negotiation problem involving the three players is below; Lemma 2 provides the payoffs that result from this negotiation (the proof is in Appendix B).

$$
\begin{aligned}
(P_{BCS}) \quad & \underset{q_S, q_C}{\text{maximize}} && q_S \cdot (q_C - c_C) \cdot (V + \delta_C - q_S - q_C) && \text{(4a)} \\
& \text{subject to} && q_S - r \geq 0, && \text{(4b)} \\
& && q_C - c_C \geq 0, && \text{(4c)} \\
& && V + \delta_C - q_S - q_C \geq 0 && \text{(4d)}
\end{aligned}
$$

Lemma 2. *In a simultaneous negotiation, the equilibrium expected payoffs U_S^T, U_B^T, and U_C^T of the seller, buyer, and consultant respectively are as follows, where $\lambda_C = \delta_C - c_C \geq 0$ is the net contribution by the consultant.*

$$
(U_S^T, U_B^T, U_C^T) = \begin{cases} \left(\frac{1}{3}(V+\lambda_C), \frac{1}{3}(V+\lambda_C), \frac{1}{3}(V+\lambda_C)\right), & \text{if } r \leq \frac{1}{3}(V+\lambda_C) \\ \left(r, \frac{1}{2}(V+\lambda_C - r), \frac{1}{2}(V+\lambda_C - r)\right), & \text{if } \frac{1}{3}(V+\lambda_C) \leq r \leq V+\lambda_C \\ (r, 0, 0), & \text{if } V+\lambda_C \leq r \end{cases}
$$

As the expected payoffs indicate, a simultaneous negotiation levels the playing field, and no one gets the advantage of being first to negotiate. Consequently, V and λ_C are weighed equally in the players' payoffs. When the outside option is small (see Lemmas 1 and 2), the seller and the buyer are able to extract more of the consultant's net contribution vis-á-vis a sequential negotiation by including the consultant in the negotiation. On the other hand, including the consultant in the negotiation changes the consultant's payoff significantly as she is now able to get a part of the data value V.

3.3. Fixed Price Consultant

So far, we have considered a specialized consultant who negotiates with the buyer to finalize the value of their services. Some consultants on the other hand, charge a fixed price for the services they provide. Let the fixed price charged by the consultant be q_C^F and the price (for the data) negotiated between the buyer and the seller be q_S. As shown earlier, the expected payoff of the buyer can be written in terms of the estimate δ_C^F of the value added by the consultant and the negotiated price of the data, q_S; this negotiation problem is below.

$$(P_{BS}^F) \quad \underset{q_S}{\text{maximize}} \quad q_S \cdot (V + \delta_C^F - q_S - q_C^F) \tag{5a}$$

$$\text{subject to} \quad q_S - r \geq 0, \tag{5b}$$

$$V + \delta_C^F - q_S - q_C^F \geq 0 \tag{5c}$$

The solution to P_{BS}^F is in Lemma 3, the proof or which is in Appendix C. $\lambda_C^F = \delta_C^F - q_C^F \geq 0$ represents the net contribution received by the buyer from the consultant after paying the fixed price.

LEMMA 3. *When the buyer hires a fixed price consultant, the equilibrium expected payoffs U_S^F and U_B^F of the seller and the buyer respectively are as below, where $\lambda_C^F = \delta_C^F - q_C^F \geq 0$.*

$$(U_S^F, U_B^F) = \begin{cases} \left(\frac{1}{2}(V+\lambda_C^F), \frac{1}{2}(V+\lambda_C^F)\right), & \text{if } r \leq \frac{1}{2}(V+\lambda_C^F) \\ \left(r, V+\lambda_C^F - r\right), & \text{if } \frac{1}{2}(V+\lambda_C^F) \leq r \leq V+\lambda_C^F \\ (r, 0), & \text{if } V+\lambda_C^F \leq r \end{cases}$$

As the buyer's decision is embedded in the seller's decision problem, the seller needs to solve the buyer's decision problems and find her own payoff given the buyer's choices. The seller can then compare this payoff with that when selling the data product and focus in on the best strategy. Next, we consider the situation where the seller also performs the role of a consultant, and sells a bundled data product.

3.4. Selling Data Product

As mentioned earlier, offering insight-driven services has created a new revenue source for data sellers who can combine data with analytic services (Banerjee et al. 2011). We now consider a situation where the seller decides to sell the data product (i.e., a combination of data and service) to the buyer, thereby eliminating the need for the buyer to hire an external consultant. Suppose the seller's services increase the value by an expected value of δ_S, at a processing cost of c_S. The seller's motivation behind this strategy is to benefit more by eliminating the need for an external consultant. The following Nash product identifies the negotiated price q_S between the seller and the buyer. Note that the implicit outside option for the seller is effectively $(r + c_S)$.

$$(P_{BS}^C) \quad \underset{q_S}{\text{maximize}} \quad (q_S - c_S) \cdot (V + \delta_S - q_S) \tag{6a}$$

$$\text{subject to} \quad q_S - c_S \geq r, \tag{6b}$$

$$V + \delta_S - q_S \geq 0 \tag{6c}$$

The solution to this problem is stated in Lemma 4; the proof is in Appendix D.

Lemma 4. *When the seller combines the data with data analytic services, the equilibrium expected payoffs U_S^C and U_B^C of the seller and the buyer respectively from the negotiation are as below, where $\lambda_S = \delta_S - c_S \geq 0$ is the net contribution by the seller (beyond the data value V).*

$$(U_S^C, U_B^C) = \begin{cases} \left(\frac{1}{2}(V + \lambda_S), \frac{1}{2}(V + \lambda_S),\right) & \textit{if } r \leq \frac{1}{2}(V + \lambda_S) \\ (r, V + \lambda_S - r,), & \textit{if } \frac{1}{2}(V + \lambda_S) \leq r \leq V + \lambda_S \\ (r, 0,), & \textit{if } V + \lambda_S \leq r \end{cases}$$

In the next section we derive the conditions under which the buyer will select the different consultant types and the negotiation format, by comparing the payoffs in Lemmas 1, 2, and 3 when the seller is selling the stand-alone data.

4. Buyer's Decision

If the seller has decided to sell only data, the buyer would first need to decide on the type of consultant to hire (type F or type N). If type N is hired, the buyer would need to decide whether to include the consultant in the data purchasing negotiation, or not. Proposition 1 provides the buyer's decision criteria associated with selecting the consultant type and the negotiation format (see Figure 2).

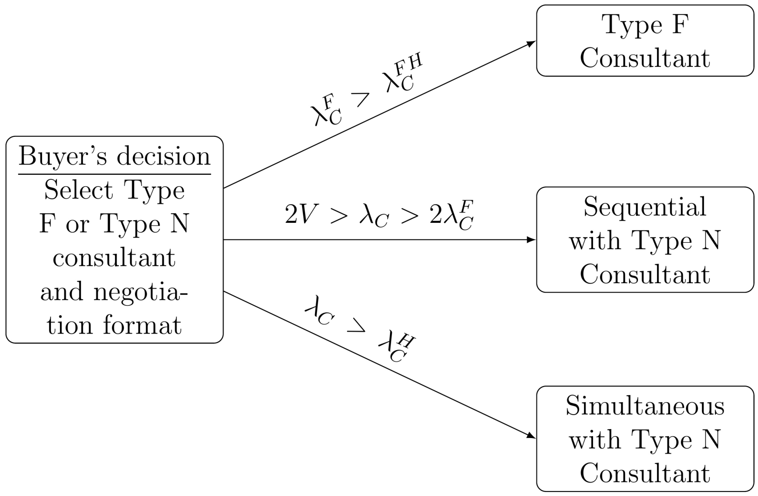


Figure 2 Buyer's Decision Criteria

PROPOSITION 1. *Given that the seller has decided to sell only data, the buyer will prefer*

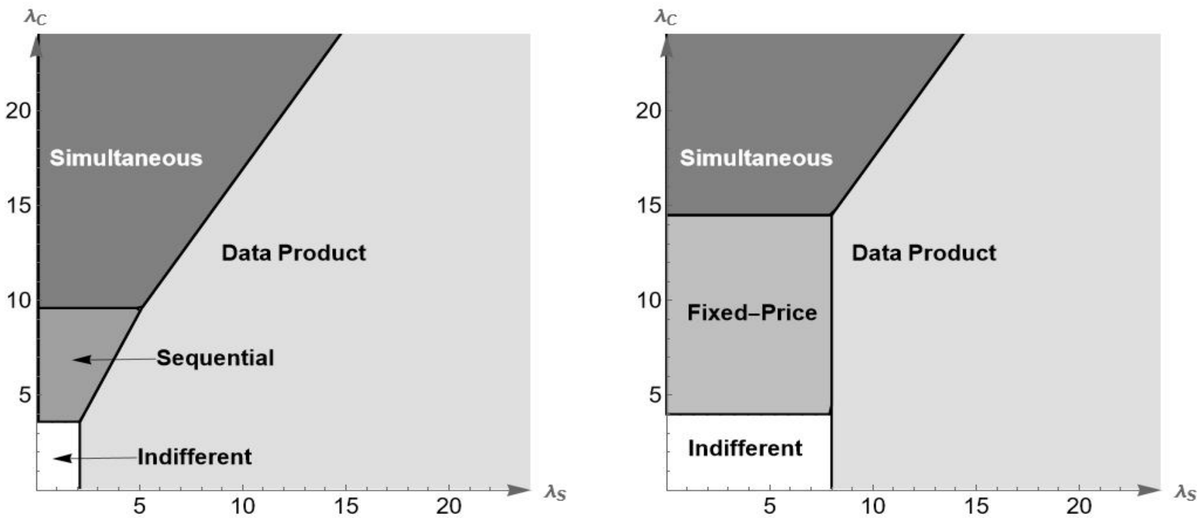
- *a sequential negotiation with a type N consultant when* $2\lambda_C^F < \lambda_C < 2V$.
- *a simultaneous negotiation with a type N consultant when* $\lambda_C > \lambda_C^H = \max\left\{2V, \frac{1}{2}(V + 3\lambda_C^F)\right\}$
- *a type F consultant when* $\lambda_C^F > \lambda_C^{FH} = \max\{\frac{1}{2}\lambda_C, \frac{1}{3}(2\lambda_C - V)\}$.

Proof: To derive the conditions for the buyer's strategy, we will use Lemma 1, Lemma 2, and Lemma 3.

The buyer would prefer a sequential negotiation to a simultaneous one when $\frac{1}{2}(V + \frac{1}{2}\lambda_C) > \frac{1}{3}(V + \lambda_C)$ i.e., when $V > \frac{1}{2}\lambda_C$. Similarly, the buyer would prefer a sequential negotiation to a to fixed-price negotiation $\frac{1}{2}(V + \frac{1}{2}\lambda_C) > \frac{1}{2}(V + \lambda_C^F)$ i.e., when $\frac{1}{2}\lambda_C > \lambda_C^F$. Therefore, given that the seller is selling only data, the buyer prefers sequential negotiation with a type N consultant to a fixed price (type F) consultant when $2\lambda_C^F < \lambda_C < 2V$.

The buyer prefers a simultaneous negotiation over a sequential one when $V < \frac{1}{2}\lambda_C$, and a simultaneous negotiation to a fixed-price negotiation when $\frac{1}{3}(V + \lambda_C) > \frac{1}{2}(V + \lambda_C^F)$ i.e., when $\lambda_C > \frac{1}{2}(V + 3\lambda_C^F)$. Therefore, given that the seller is selling only data, the buyer will prefer a simultaneous negotiation with a type N consultant when $\max\left\{2V, \frac{1}{2}(V + 3\lambda_C^F)\right\} < \lambda_C$.

It follows from this analysis that the buyer is better-off hiring a type F when $\lambda_C^F > \max\{\frac{1}{2}\lambda_C, \frac{1}{3}(2\lambda_C - V)\}$. ■

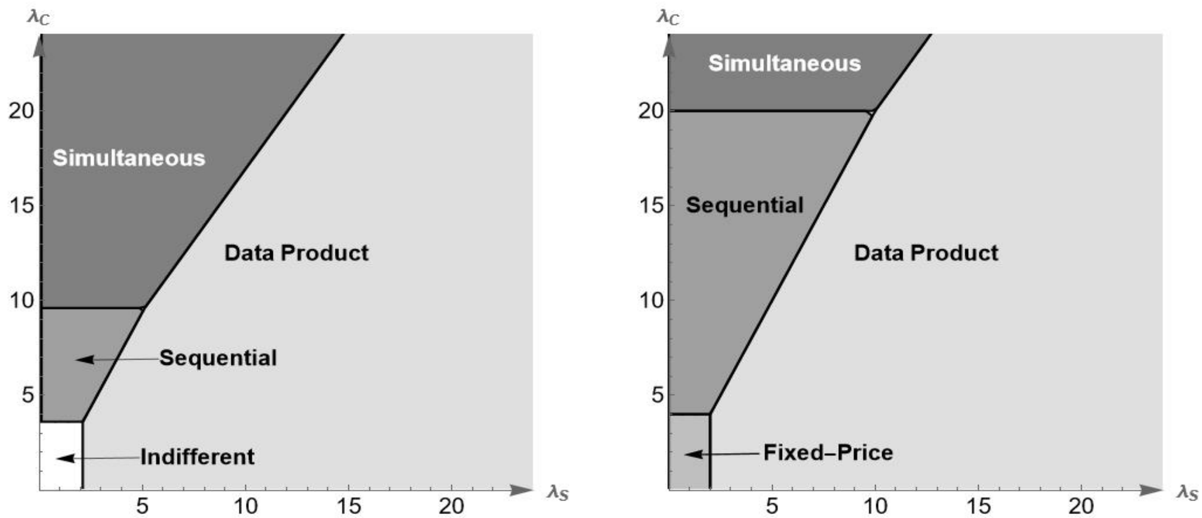


(a) When $\lambda_C^F = 2, V = 5, r = 3$ (b) When $\lambda_C^F = 8, V = 5, r = 3$

Figure 3 Effect of λ_C^F on Seller's and Buyer's Decisions

Figures 3 and 4 show the regions where the buyer and the seller make decisions (the seller's decision is discussed in the next section). Let us consider the buyer's choice between consultant types N and F. If the net contribution received by the buyer from the fixed-price consultant is very small ($\lambda_C^F < \max\left\{\frac{\lambda_C}{2}, \frac{2\lambda_C - V}{3}\right\}$) the buyer would prefer to hire the type N consultant (see Figure 3(a)). Interestingly, the buyer's decision reverses when λ_C^F becomes moderate relative to λ_C (while still not very high), and r is small (see Figure 3(b)). This is the situation where a fixed-price consultant benefits the buyer – even though the net contribution (λ_C) of the type N consultant is somewhat higher than that of the fixed-price consultant (λ_C^F), it is shared among all three players. Therefore, the buyer receives only a small part of λ_C – less than the added contribution from the fixed-price option.

The choice between simultaneous and sequential negotiations with a type N consultant is a direct result of the relative importance of V and λ_C in the buyer's equilibrium payoff. Let us first consider the case when both λ_C^F and r are relatively small (see Figures 4(a) and 4(b)). As mentioned earlier, the consultant cannot claim any part of the data value V in the sequential negotiation; the consultant's payoff only depends on the net contribution λ_C. This results in the seller and the buyer sharing V between each other when the outside option is small. On the other hand, the consultant can take advantage of the fact that the data has already been purchased and that the buyer needs the



(a) When $\lambda_C^F = 2, V = 5, r = 3$ (b) When $\lambda_C^F = 2, V = 10, r = 3$

Figure 4 Effect of V and λ_C on Seller's and Buyer's Decisions

consultant for insights. In a sequential negotiation, the consultant ends up taking the larger share of λ_C, while the buyer and the seller receives a smaller portion. As a result, the equilibrium payoffs of the seller and the buyer have a larger share of V than of λ_C in the sequential negotiation, as underscored by Lemma 1. In contrast, the simultaneous negotiation provides a level-playing field for all three players, and the combined value of the data and services ($V + \lambda_C$) is shared equally among them when r is small. Consequently, the buyer's payoff consists of a relatively large proportion of λ_C in the simultaneous negotiation compared to the sequential negotiation when r is small. So unless λ_C is substantially larger than V, the buyer will select a sequential negotiation over a simultaneous one when r is small. In summary, if the value of the data is high relative to the added value from the consultant, the buyer will not want to share it with a type N consultant and opt for a sequential negotiation. If the additional contribution from the type N consultant is significant compared to the value of the data, it would make sense to include the consultant in the data purchasing negotiation with the seller in order to extract more of the consultant's net contribution.

In Lemmas 1–3, we observe that the payoffs for the seller and the buyer are the same when the outside option is small. This implies that the buyer's decisions of selecting the consultant type and the negotiation format are similarly preferred by the seller. This is due to the fact that their incentives are aligned. Since the payoffs of both the seller and the buyer increase with λ_C and λ_C^F when stand-alone data is sold, it becomes imperative for both of them to hire the consultant who provides the highest net contribution. On the contrary, the incentives of the buyer and the consultant always are at odds. When V is large, the buyer would prefer not to share it with the consultant in a simultaneous negotiation, preferring a sequential negotiation instead (thereby, keeping the consultant away from the data buying process). Conversely, the consultant would prefer a simultaneous negotiation when V is large as they can get a share of V (see Lemma 2), which is not possible in a sequential negotiation.

5. Seller's Decision

The seller is interested in knowing when selling the data product is preferable to selling just the raw data. To answer this question, the seller can leverage Proposition 1 and identify conditions under which the buyer will hire a type N and a type F consultant. The seller can then find their own payoffs when the buyer decides among the sequential (Lemma 1), simultaneous (Lemma 2), or fixed-price (Lemma 3) negotiations. The seller's payoffs from selling only data can then be compared with the payoff when they also act as a consultant (Lemma 4). Proposition 2 provides a threshold value of λ_S^H for the seller beyond which it is optimal to sell the data product (see Figure 5).

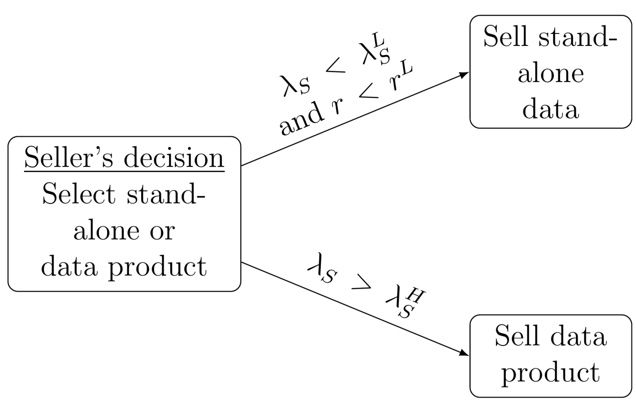


Figure 5 Seller's Decision Criteria

Proposition 2. *The seller will*

- *prefer selling the data product when* $\lambda_S > \lambda_S^H = \max\left\{\frac{\lambda_C}{2}, \lambda_C^F, \frac{2\lambda_C - V}{3}, 2r - V\right\}$.
- *prefer selling only the data when* $\lambda_S < \lambda_S^L = \max\left\{\frac{\lambda_C}{2}, \lambda_C^F, \frac{2\lambda_C - V}{3}\right\}$, *and* $r < r^L = \max\left\{\frac{V+\lambda_C}{3}, \frac{V+\frac{1}{2}\lambda_C}{2}, \frac{V+\lambda_C^F}{2}\right\}$.

- *be indifferent otherwise.*

Proof: Note that the seller will prefer selling either stand-alone data or a data product when her equilibrium payoff from the negotiations is greater than the outside option r. This happens when r is small, that is, $r < \frac{1}{2}(V + \lambda_C^F)$ in fixed-price negotiation, $r < \frac{1}{2}(V + \frac{1}{2}\lambda_C)$ in sequential negotiation, $r < \frac{1}{3}(V + \lambda_C)$ in simultaneous negotiation, and $r < \frac{1}{2}(V + \lambda_S)$ when data product is sold. Since the buyer's decision is embedded in the seller's decision, the seller will first identify the buyer's best strategy (consultant type and negotiation format) when only data is sold. Based on the buyer's best strategy, the seller will find her own equilibrium payoff when only data is sold and compare that with her payoff when the data product is sold. Interestingly, the equilibrium payoffs of the seller and the buyer are the same when r is small (refer to the first condition in Lemma 1, Lemma 2, and Lemma 3). Hence, the strategy preferred by the buyer will also be preferred by the seller. Therefore, the seller only needs to compare her equilibrium payoffs in the sequential, simultaneous, and fixed-price negotiations with her payoff when the data product is sold.

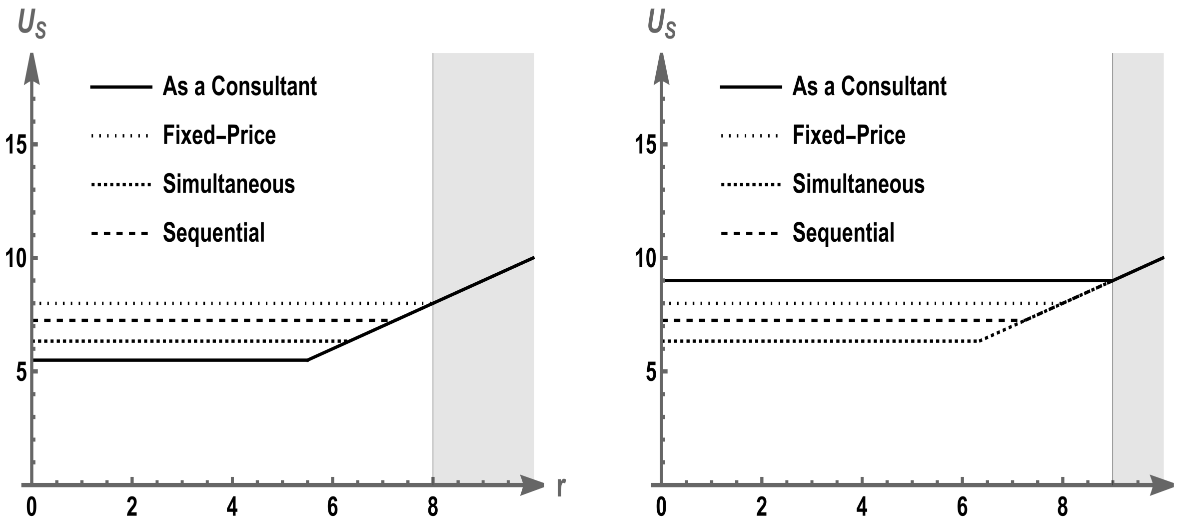
Let us consider the case when $r < \frac{1}{2}(V + \frac{1}{2}\lambda_C)$ and $r < \frac{1}{2}(V + \lambda_S)$. From Lemmas 1 and 4, the seller's equilibrium payoff from selling the data product is more than that from the sequential negotiation when $\frac{1}{2}(V + \frac{1}{2}\lambda_C) < \frac{1}{2}(V + \lambda_S)$ i.e., $\frac{1}{2}\lambda_C < \lambda_S$. Therefore, the seller will prefer to sell the data product compared to a sequential negotiation with a type N consultant when $\max\left\{\frac{1}{2}\lambda_C, 2r - V\right\} < \lambda_S$.

We next derive the condition when the seller prefers to sell the data product over a simultaneous negotiation. This happens in the region $r < \frac{1}{3}(V + \lambda_C)$ and $r < \frac{1}{2}(V + \lambda_S)$. From Lemmas 2 and 4, the seller's equilibrium payoff from selling the data product is more than that from the simultaneous negotiation when $\frac{1}{3}(V + \lambda_C) < \frac{1}{2}(V + \lambda_S)$ i.e., when $\lambda_S > \frac{1}{3}(2\lambda_C - V)$. Therefore, the seller will sell the data product (compared to a simultaneous negotiation with a type N consultant) when $\max\left\{\frac{1}{3}(2\lambda_C - V), 2r - V\right\} < \lambda_S$.

The region of r under which the seller prefers selling the data product over a type F consultant is given by $r < \frac{1}{2}(V + \lambda_C^F)$ and $r < \frac{1}{2}(V + \lambda_S)$. From Lemmas 3 and 4, the seller's equilibrium payoff from selling the data product is more than that from the fixed-price consultant when $\frac{1}{2}(V + \lambda_C^F) <$

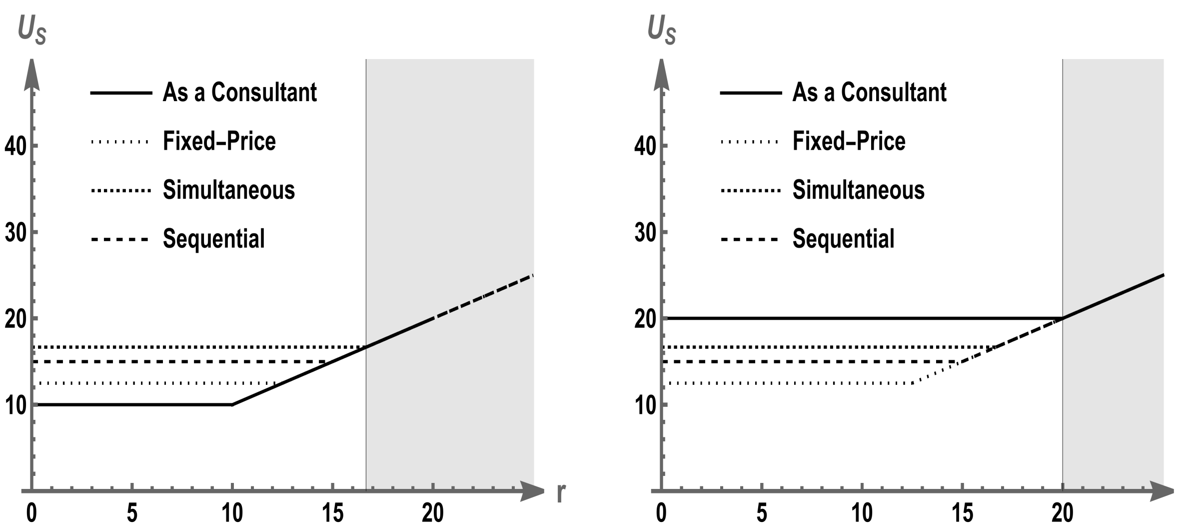
$\frac{1}{2}(V+\lambda_S)$ i.e., when $\lambda_C^F < \lambda_S$. Therefore, the seller will prefer to sell the data product compared to selling stand-alone data with the buyer hiring a type F consultant when $\max\{\lambda_C^F, 2r-V\} < \lambda_S$.

Combining all the criteria above, we observe that the seller will prefer selling the data product when $\lambda_S > \max\{\frac{1}{2}\lambda_C, \lambda_C^F, \frac{1}{3}(2\lambda_C - V), 2r-V\}$. Conversely, the seller will prefer selling stand-alone data when $\lambda_S < \max\{\frac{1}{2}\lambda_C, \lambda_C^F, \frac{1}{3}(2\lambda_C - V)\}$, and $r < \max\{\frac{1}{3}(V+\lambda_C), \frac{1}{2}(V+\frac{1}{2}\lambda_C), \frac{1}{2}(V+\lambda_C^F)\}$. ■



(a) When $\lambda_S = 1, \lambda_C = 9, \lambda_C^F = 6, V = 10$ (b) When $\lambda_S = 8, \lambda_C = 9, \lambda_C^F = 6, V = 10$

Figure 6 Comparison of Seller's Payoff with Relatively Large λ_C^F



(a) When $\lambda_S = 10, \lambda_C = 40, \lambda_C^F = 15, V = 10$ (b) When $\lambda_S = 30, \lambda_C = 40, \lambda_C^F = 15, V = 10$

Figure 7 Comparison of Seller's Payoff with Relatively Large λ_C

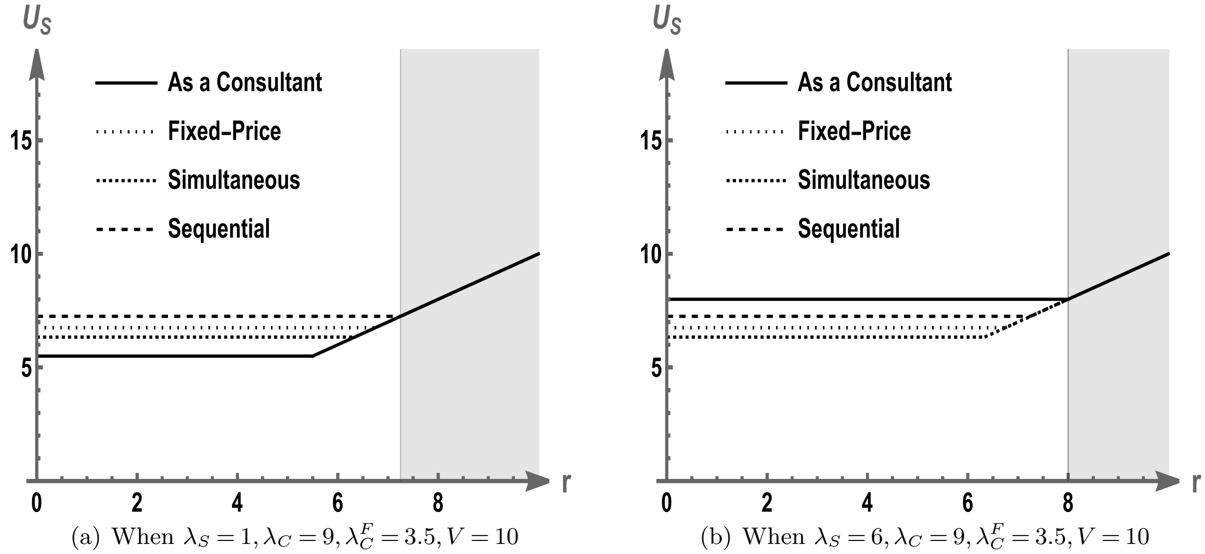
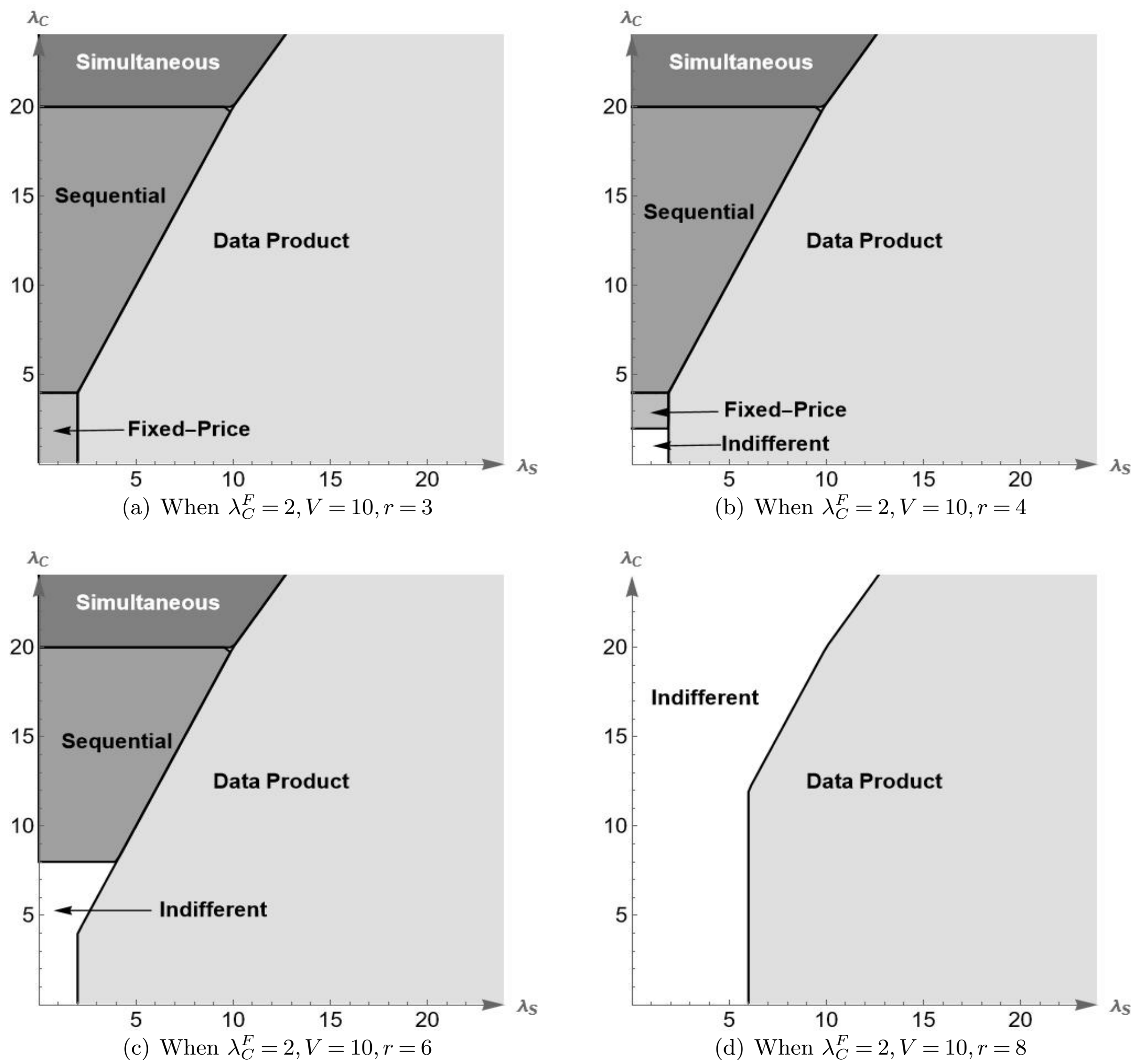


Figure 8 **Comparison of Seller's Payoff with Relatively Large** V

Figures 6, 7, and 8 plot the seller's equilibrium payoff as a consultant, and compares it with the payoffs from fixed-price, simultaneous, and sequential negotiations respectively. The seller will sell data bundled with the service when the added net contribution ($\lambda_S = \delta_S - c_S$) from the service is high relative to the consultant's net contribution ($\lambda_C = \delta_C - c_C$ and $\lambda_C^F = \delta_C^F - q_C^F$), as it ensures a better consultancy capability. An interesting observation is that the seller will prefer selling the data product even when their added contribution is less than that of a type N consultant but good enough to compete. This is illustrated in Figures 6(b), 7(b), and 8(b) when the outside option is small. By providing data analytic services, the seller is able to eliminate the need for a third-party consultant entirely. Thus, even though the net contribution of the seller is lower than that of the type N consultant, the total value is divided only between the seller and buyer (as there is no consultant), which benefits both. In summary, the seller will not sell stand-alone data unless the type N consultant adds significant value through their services. This is also supported by the fact that the threshold value λ_S^H increases with λ_C.

If the added value from the seller's consultancy services is not good enough to compete, and if she does not have a good outside option to fall back on, she will abandon the idea of selling the data product and instead will try to extract a part of the consultant's added contribution through the sequential or simultaneous negotiation process as observed in Figures 7(a) and 8(a). Lemmas 3 and 4 indicate that the seller has the same impact on her payoff as a consultant compared to the case when the buyer hires a type F consultant. Hence, to beat a type F consultant, the seller either needs to be a cost leader ($c_C < q_C^F$) when their contributions are the same, or provide a higher added contribution than what the type F consultant does when $c_C = q_C^F$.



(a) When $\lambda_C^F = 2, V = 10, r = 3$
(b) When $\lambda_C^F = 2, V = 10, r = 4$
(c) When $\lambda_C^F = 2, V = 10, r = 6$
(d) When $\lambda_C^F = 2, V = 10, r = 8$

Figure 9 Effect or r on Seller's and Buyer's Decisions

As the outside option r increases, the region of selling stand-alone data shrinks as the seller becomes increasingly indifferent (see Figures 9(a)–9(d)). In this indifference region (the shaded area in Figures 6–8) the seller receives the same payoff as the outside option, and hence, there is no gain from selling the data to the buyer. A higher outside option eventually leads to an unsuccessful negotiation. On the other hand, as λ_C increases, the region associated with selling just the stand-alone data expands since the seller can now benefit from extracting a part of λ_C via negotiation (see Figures 3, 4, and 9). Table 2 lists the players' decisions at different parameter levels and thresholds respectively.

Player	**Criteria**	**Decision**
Seller	Moderate to high λ_S and low r	Sell data-product
	Low λ_S and low r	Sell stand-alone data
	Moderate to high r	Indifferent
Buyer	Moderate to high λ_C^F	Hire Type F
	Low λ_C^F and high λ_C relative to V	Simultaneous negotiation with Type N
	Low λ_C^F and low λ_C relative to V	Sequential negotiation with Type N

Table 2 Players' Decisions at Different Levels of Parameters

6. Type N and F Consultants with Identical Contributions

There could be situations where both consultant types provide similar value. Buyers need to analyze this situation and obtain the best strategy when the added contribution (δ) is the same for both type N and type F consultants. Proposition 3 provides the buyer's decision in this situation.

PROPOSITION 3. *Given that seller has decided to sell only stand-alone data and both type N and type F consultants add the same value (δ) by their services, buyer will prefer*

- *a sequential negotiation with type N consultant when* $\delta < \min\{2q_C^F - c_C, 2V + c_C\}$
- *a simultaneous negotiation with type N consultant when* $2V + c_C < \delta < 3q_C^F - 2c_C - V$
- *hiring type F consultant when* $\delta > \max\{2q_C^F - c_C, 3q_C^F - 2c_C - V\}$.

Proof: To verify this, we use Proposition 1 and set $\delta_C = \delta_C^F = \delta$. Given that the seller has decided to sell only stand-alone data, the buyer will prefer a sequential negotiation with a type N consultant when $\lambda_C^F < \frac{1}{2}\lambda_C < V$. Replacing $\lambda_C^F = \delta - q_C^F$ and $\lambda_C = \delta - c_C$, we get $\delta < 2q_C^F - c_C$ and $\delta < 2V + c_C$ respectively. Therefore, $\delta < \min\{2q_C^F - c_C, 2V + c_C\}$ satisfies the condition for the buyer to select a sequential negotiation. The buyer will prefer a simultaneous negotiation with a type N consultant when $\lambda_C > 2V$ and $\lambda_C > \frac{1}{2}(V + 3\lambda_C^F)$. These inequalities give the following conditions in favor of a simultaneous negotiation: $2V + c_C < \delta < 3q_C^F - 2c_C - V$. The buyer will hire a type F consultant when $\lambda_C^F > \frac{1}{2}\lambda_C$ and $\lambda_C^F > \frac{1}{3}(2\lambda_C - V)$. These imply that a fixed-price consultant would be preferred if $\delta > \max\{2q_C^F - c_C, 3q_C^F - 2c_C - V\}$. ∎

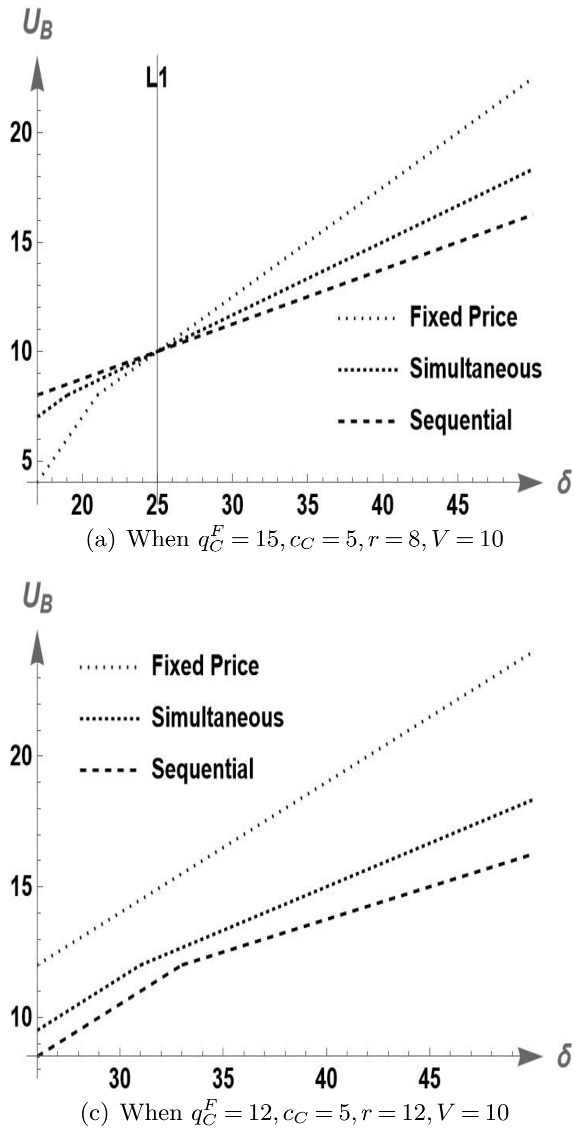


Figure 10 Comparison of Buyer's Payoff when δ is same for Type F and Type N Consultants

Proposition 3 is best understood through Figure 10 which plots the buyer's equilibrium payoffs against δ. As already noted, the buyer will prefer a sequential negotiation when the net contribution $(\delta - c_C)$ is relatively small compared to V (refer to the left of line L1 in Figures 10(a) and 10(b)). When δ is very large relative to the fixed-price q_C^F, the buyer will opt not to share this large added value with the consultant. The buyer would prefer to pay the fixed-price to the type F consultant and divide the large δ between the seller and herself as seen in Figures 10(b) and 10(d) (beyond line L2). When δ is moderately larger than V, the decision of choosing between a type F consultant and a simultaneous negotiation with a type N consultant is determined by the fixed-price q_C^F and cost c_C. If q_C^F is large ($q_C^F > V + c_C$), the buyer will prefer not to pay the high fixed price and have a simultaneous negotiation instead, as the net value $(V + \delta - c_C)$ in the simultaneous negotiation is much more than that in the fixed-price option $(V + \delta - q_C^F)$ (refer to the plot between lines L1 and L2 in Figure 10(b) and left of line L2 in Figure 10(d)). Even though the buyer only receives a third of the net value in the simultaneous negotiation, it is preferable to paying a high fixed price.

7. Conclusion

This paper examines the nexus of relationships between data sellers, data buyers, and analytic service providers (consultants) when data and complementary analytic services are purchased. Because our context emphasizes the exclusive selling of proprietary and unique data (rather than general purpose data that can be sold to many buyers), announcing a fixed price for the data is not a viable option for the seller. Thus we use a Nash bargaining framework where the negotiations always involve the data seller and data buyer and could sometimes involve a consultant. We study the outcomes from the perspectives of both the data seller and the data buyer. Both these perspectives are important if the market for data and analytic services is to prosper. In a world where the monetizing of data is becoming common, data sellers could reduce investment in data gathering and cleaning activities. On the other hand, if data buyers cannot efficiently buy data and obtain complementary analytic services, there will be less demand for data and analytic services in the market. By maximizing the product of the payoffs for sellers and buyers (and the payoff for the consultant if applicable), the Nash bargaining framework tries to ensure that efficient outcomes for all concerned parties are achieved.

We find that a relatively small policy shift (whether the consultant is engaged simultaneously or sequentially) can lead to significantly different outcomes for the data buyer. Another important decision for the data buyer is whether to hire a consultant who announces a fixed price for analytic services or hire one whose price for services is subject to the outcome of a bargaining process. A fixed price consultant is preferable when the net contribution from this consultant is relatively high. On the other hand, when it is better for the buyer to hire a consultant who prefers to negotiate the price charged for analytic services, it matters whether such a consultant is brought into the picture simultaneously (while the price of the data is being negotiated) or sequentially (after the price of the data has been negotiated). The buyer would prefer to simultaneously negotiate the price of the data and analytic services if the consultant is of relatively high capability. Otherwise, it would be better for the buyer to separately bargain with the data seller on the price of data and with the consultant on the price of analytic services. From the perspective of the data seller, we determine when bundling data with complementary analytic services can improve outcomes. The data seller has a natural advantage over a specialized consultant concerning the provision of analytic services. Even if the data seller cannot match the consultant's capabilities, bundling data with analytic services can be a superior option for the data seller. Thus data sellers should strive to create in-house analytic service capabilities if they wish to fully monetize their data.

Our study is not without limitations. While we allow for uncertainty in the valuation of the data and the value added by the consultant, we assume that all the information is symmetric and known to all parties. Studies have shown that buyers often underestimate the value of information goods (like data) in the face of uncertainty. While our model does account for uncertainty, we do not consider underestimation in this paper. It would be interesting to identify conditions when it might benefit the seller to provide a demonstration to better signal the value of the data to the buyer. The involvement of the consultant presents an interesting trade-off – to some extent, the consultant can play the role of the demonstration by making the buyer better informed, and thereby reducing the buyer's natural tendency to underestimate. In some situations it is possible that these roles to reduce underestimation (played by the demonstration and the consultant) could complement each other, but in other circumstances substitute one another. Future work could investigate the joint roles of the consultant and a demonstration in the presence of underestimation.

References

Agarwal A, Dahleh M, Sarkar T (2019) A Marketplace for Data: An Algorithmic Solution. *Proceedings of the 2019 ACM Conference on Economics and Computation*, 701–726.

Babaioff M, Kleinberg R, Paes Leme R (2012) Optimal Mechanisms for Selling Information. *Proceedings of the 13th ACM Conference on Electronic Commerce*, 92–109.

Bakos Y, Brynjolfsson E (1999) Bundling Information Goods: Pricing, Profits, and Efficiency. *Management Science* 45(12):1613–1630.

Bakos Y, Brynjolfsson E (2000) Bundling and competition on the internet. *Marketing Science* 19(1):63–82.

Banerjee J Sand Bolze, McNamara J, O'Reilly K (2011) How Big Data can Fuel Bigger Growth. *Accenture Outlook* (3):28–35.

Belissent J (2020) Data Commercialization: A CIO's Guide To Taking Data To Market. *Forrester* March 27, 2020. Available at `https://datadrivenenterprise.cio.com/data-commercialization-a-cios-guide-to-taking-data-to-market/`.

Bennett E (1997) Multilateral bargaining problems. *Games and Economic Behavior* 19(2):151–179.

Bennett E, Houba H, et al. (1995) *Odd man out: bargaining among three players* (Tinbergen Institute).

Bergemann D, Bonatti A (2015) Selling Cookies. *American Economic Journal: Microeconomics* 7(3):259–294.

Bergemann D, Bonatti A, Gan T (2022) The Economics of Social Data. *The RAND Journal of Economics* 53(2):263–296.

Bhargava H, Csapó G, Müller R (2020) On Optimal Auctions for Mixing Exclusive and Shared Matching in Platforms. *Management Science* 66(6):2653–2676.

Bimpikis K, Crapis D, Tahbaz-Salehi A (2019) Information Sale and Competition. *Management Science* 65(6):2646–2664.

Binmore K (1985) Bargaining and Coalitions. Roth A, ed., *Game-theoretic Models of Bargaining*, chapter 13, 269–304 (Cambridge University Press).

Binmore K, Shared A, Sutton J (1989) An outside option experiment. *The Quarterly Journal of Economics* 104(4):753–770.

Burguet R, Caminal R (2011) Simultaneous nash bargaining with consistent beliefs .

Compte O, Jehiel P (2002) On the Role of Outside Options in Bargaining with Obstinate Parties. *Econometrica* 70(4):1477–1517.

Deal J (2018) Hiring a Data Analytics Consultant. *Elder Research, Inc.* Oct. 5, 2018. Available at `https://www.elderresearch.com/blog/hiring-a-data-analytics-consultant/`.

Deichmann J, Heineke T Kand Reinbacher, Wee D (2016) Creating a Successful Internet of Things Data Marketplace. *McKinsey Insights* October 7, 2016. Available at `https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/creating-a-successful-internet-of-things-data-marketplace#`.

Deloitte (2018) Data Takes Daimler Trucks Asia on an Insights-Driven Journey. *The Wall Street Journal* Apr 17, 2018.

Ferme W (1987) What Could a Client Reasonably Expect a Consultant to do When He is Engaged to Assist in the Purchase of a Manufacturing System. *Australian Computer Journal* 19(4):196–202.

Gable G (1991) Consultant engagement for computer system selection: a pro-active client role in small businesses. *Information & Management* 20(2):83–93.

Gandhi S, Thota B, Kuchembuck R, Swartz J (2018) Demystifying Data Monetization. *MIT Sloan Management Review* .

Gans JS, Hsu DH, Stern S (2008) The impact of uncertain intellectual property rights on the market for ideas: Evidence from patent grant delays. *Management Science* 54(5):982–997.

Geng X, Stinchcombe M, Whinston A (2005) Bundling Information Goods of Decreasing Value. *Management Science* 51(4):662–667.

Ghosh A, Roth A (2011) Selling privacy at auction. *Proceedings of the 12th ACM conference on Electronic commerce*, 199–208.

Gkatzelis V, Aperjis C, Huberman B (2015) Pricing Private Data. *Electronic Markets* 25(2):109–123.

Henry D (2017) How American Airlines is Innovating Customer Travel with IBM Cloud. *IBM THINK Blog* June 27, 2017.

Holst A (2021) Amount of Data Created, Consumed, and Stored 2010-2025. *Statista* June 7, 2021. Available at `https://www.statista.com/statistics/871513/worldwide-data-created`.

Jung K, Lee J, Park K, Park S (2019) Privata: differentially private data market framework using negotiation-based pricing mechanism. *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 2897–2900.

Koutris P, Upadhyaya P, Balazinska M, Howe B, Suciu D (2015) Query-based data pricing. *Journal of the ACM (JACM)* 62(5):1–44.

Kulkarni R (2019) Big Data Goes Big. *Statista* February 7, 2019. Available at `https://www.forbes.com/sites/rkulkarni/2019/02/07/big-data-goes-big/`.

Kushal A, Moorthy S, Kumar V (2012) Pricing for Data Markets Working Paper. Available at `https://www.semanticscholar.org/paper/Pricing-for-Data-Markets-Kushal-Moorthy/93c70f67a329279afe66a753252f546642a0e1aa`.

Li C, Giampapa J, Sycara K (2006) Bilateral negotiation decisions with uncertain dynamic outside options. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 36(1):31–44.

Lotame (2020) How to Monetize Your Data. Published: January 13, 2020; Last modified: March 21, 2022. Available at `https://www.lotame.com/how-to-monetize-your-data/`.

Mehta S, Dawande M, Janakiraman G, Mookerjee V (2021) How to Sell a Data Set? Pricing Policies for Data Monetization. *Information Systems Research* 32(4):1281–1297.

Montgomery A (1987) The Australian Computer Journal: a twentieth anniversery editorial: Twenty Deadly Sins Against Professionalism and Progress in Australian Computing. *Australian Computer Journal* 19(4):190–195.

Najjar M, Kettinger W (2013) Data Monetization: Lessons from a Retailer's Journey. *MIS Quarterly Executive* 12(4).

Nash J (1950) The Bargaining Problem. *Econometrica: Journal of the Econometric Society* 155–162.

Ponsatí C, Sákovics J (1998) Rubinstein Bargaining with Two-sided Outside Options. *Economic Theory* 11(3):667–672.

Radhakrishnan D, Heizenberg J, Duncan A, Jones L (2020) How to Monetize Data Assets With Your Data and Analytics Service Provider. *Gartner* July 16, 2020. Available at `https://www.gartner.com/en/documents/3987625`.

Ray J, Menon S, Mookerjee V (2020) Bargaining over Data: When Does Making the Buyer More Informed Help? *Information Systems Research* 31(1):1–15.

Rochford SC (1984) Symmetrically pairwise-bargained allocations in an assignment market. *Journal of Economic Theory* 34(2):262–281.

Rubinstein A, Safra Z, Thomson W (1992) On the interpretation of the Nash bargaining solution and its extension to non-expected utility preferences. *Econometrica: Journal of the Econometric Society* 1171–1186.

Salinger M (1995) A Graphical Analysis of Bundling. *Journal of Business* 85–98.

Sundararajan A (2004) Nonlinear Pricing of Information Goods. *Management Science* 50(12):1660–1673.

Walden EA (2005) Intellectual property rights and cannibalization in information technology outsourcing contracts. *MIS Quarterly* 699–720.

Wixom B, Ross J (2017) How to Monetize Your Data. *MIT Sloan Management Review* 58(3).

Wu S, Hitt L, Chen P, Anandalingam G (2008) Customized Bundle Pricing for Information Goods: A Nonlinear Mixed-Integer Programming Approach. *Management Science* 54(3):608–622.

Appendix

A. Proof of Lemma 1

Lemma 1. *In a sequential negotiation, the equilibrium expected payoffs U_S^S, U_B^S, and U_C^S of the seller, buyer, and consultant respectively are as follows where $\lambda_C = \delta_C - c_C \geq 0$ is the net contribution by the consultant,*

$$(U_S^S, U_B^S, U_C^S) = \begin{cases} \left(\frac{1}{2}(V + \frac{1}{2}\lambda_C), \frac{1}{2}(V + \frac{1}{2}\lambda_C), \frac{1}{2}\lambda_C\right), & \textit{if } r \leq \frac{1}{2}(V + \frac{1}{2}\lambda_C) \\ \left(r, V + \frac{1}{2}\lambda_C - r, \frac{1}{2}\lambda_C\right), & \textit{if } \frac{1}{2}(V + \frac{1}{2}\lambda_C) \leq r \leq V + \frac{1}{2}\lambda_C \\ (r, 0, 0), & \textit{if } V + \frac{1}{2}\lambda_C \leq r \end{cases}$$

Proof: We will first solve the negotiation problem (2a) between buyer and consultant. We consider the Karush-Kuhn-Tucker (KKT) conditions and write the Lagrangian as $\mathcal{L}(q_C, \lambda_1, \lambda_2) = (q_C - c_C) \cdot (\delta_C - q_C) - \lambda_1(q_C - c_C) - \lambda_2(\delta_C - q_C)$. The KKT conditions are represented by the following equations, along with the non-negativity Lagrangian multipliers $\lambda_1 \geq 0$, and $\lambda_2 \geq 0$:

$$\begin{aligned} \tfrac{\partial \mathcal{L}}{\partial q_C} : \delta_C - 2q_C + c_C - \lambda_1 + \lambda_2 &= 0 \quad (7) \\ \lambda_1(q_C - c_C) &= 0 \\ \lambda_2(\delta_C - q_C) &= 0 \\ q_C - c_C &\geq 0 \\ \delta_C - q_C &\geq 0 \end{aligned}$$

Since $q_C = c_C$ and $q_C = \delta_C$ makes the objective function 0, players can do better by setting $\delta_C > q_C > c_C$. This implies that $\lambda_1 = 0$ and $\lambda_2 = 0$. From (7) we get $q_C^* = \frac{\delta_C + c_C}{2}$ which makes the objective function $\left(\frac{\delta_C - c_C}{2}\right)^2 > 0$. The equilibrium expected payoffs of the buyer and the consultant are $(U_B, U_C) = \left(\frac{\delta_C - c_C}{2}, \frac{\delta_C - c_C}{2}\right)$ which is feasible only when $c_C \leq \delta_C$.

We will next solve the negotiation problem (3a) between buyer and seller. We formulate the Lagrangian of the problem as $\mathcal{L}(q_S, \lambda_1, \lambda_2) = q_S(V' - q_S) - \lambda_1(q_S - r) - \lambda_2(V' - q_S)$ where $V' = V + \delta_C - q_C^*$. The KKT conditions are represented by the following equations, along with the non-negativity constraints $\lambda_1 \geq 0$, and $\lambda_2 \geq 0$:

$$\frac{\partial \mathcal{L}}{\partial q_S}: V' - 2q_S - \lambda_1 + \lambda_2 = 0 \quad (8)$$

$$\lambda_1(q_S - r) = 0$$

$$\lambda_2(V' - q_S) = 0 \quad (9)$$

$$q_S - r \geq 0$$

$$V' - q_S \geq 0$$

A.1. Case I: $q_S \neq r$

This implies that $\lambda_1 = 0$ and equation (8) will give $\lambda_2 = 2q_S - V'$. From (9) we get $(2q_S - V')(q_S - V') = 0$ which gives $q_S = V', \frac{V'}{2}$. The objective is 0 when $q_S = V'$, and $\left(\frac{V'}{2}\right)^2$ when $q_S = \frac{V'}{2}$; therefore, the optimal solution when $q_S \neq r$ is $q_S = \frac{V'}{2}$. Furthermore, $(U_S, U_B) = \left(\frac{V'}{2}, \frac{V'}{2}\right)$ is feasible only when $r \leq \frac{V'}{2}$.

A.2. Case II: $q_S = r$

This implies that $U_B(q_S) = V' - r$. For a feasible solution, we need $V' \geq r$.

Since both $q_S = r$ and $q_S = \frac{V'}{2}$ are feasible solutions in the range $r < \frac{V'}{2}$, we need to compare their objective values $\left(\frac{V'}{2}\right)^2$ and $r(V' - r)$ to find the global optimum. We compare the arithmetic mean and geometric mean of r and $V' - r$ i.e. $\frac{1}{2}(r + (V' - r)) \geq \sqrt{r(V' - r)}$ which implies $\left(\frac{V'}{2}\right)^2 \geq r(V' - r)$. Therefore, $\left(\frac{V'}{2}, \frac{V'}{2}\right)$ is the global solution when $r \leq \frac{V'}{2}$ and $(r, V' - r)$ is the global solution in the range $\frac{V'}{2} \leq r \leq V'$. For $r \geq V'$, none of the prices satisfy all the conditions, and therefore, there will be no agreement – the disagreement outcome $(r, 0)$ will result. ∎

B. Proof of Lemma 2

LEMMA 2. *In a simultaneous negotiation, the equilibrium expected payoffs U_S^T, U_B^T, and U_C^T of the seller, buyer, and consultant respectively are as follows where $\lambda_C = \delta_C - c_C \geq 0$ is the net contribution by the consultant,*

$$(U_S^T, U_B^T, U_C^T) = \begin{cases} \left(\frac{1}{3}(V + \lambda_C), \frac{1}{3}(V + \lambda_C), \frac{1}{3}(V + \lambda_C)\right), & \text{if } r \leq \frac{1}{3}(V + \lambda_C) \\ \left(r, \frac{1}{2}(V + \lambda_C - r), \frac{1}{2}(V + \lambda_C - r)\right), & \text{if } \frac{1}{3}(V + \lambda_C) \leq r \leq V + \lambda_C \\ (r, 0, 0), & \text{if } V + \lambda_C \leq r \end{cases}$$

Proof: The three player negotiation problem is given by (4a). The Lagrangian of the problem is $\mathcal{L}(q_S, q_C, \lambda_1, \lambda_2, \lambda_3) = q_S(q_C - c_C)(V + \delta_C - q_S - q_C) - \lambda_1(q_S - r) - \lambda_2(q_C - c_C) - \lambda_3(V + \delta_C - q_S - q_C)$. The KKT conditions are as follows, along with the non-negativity constraints $\lambda_1 \geq 0$, $\lambda_2 \geq 0$, and $\lambda_3 \geq 0$:

$$\frac{\partial \mathcal{L}}{\partial q_S}: \; (q_C - c_C)(V + \delta_C - 2q_S - q_C) - \lambda_1 + \lambda_3 = 0 \quad (10)$$

$$\frac{\partial \mathcal{L}}{\partial q_C}: \; q_S(V + \delta_C - q_S - 2q_C + c_C) - \lambda_2 + \lambda_3 = 0 \quad (11)$$

$$\lambda_1(q_S - r) = 0$$

$$\lambda_2(q_C - c_C) = 0$$

$$\lambda_3(V + \delta_C - q_S - q_C) = 0$$

$$q_S - r \geq 0$$

$$q_C - c_C \geq 0$$

$$V + \delta_C - q_S - q_C \geq 0$$

Since $q_C - c_C = 0$ and $V + \delta_C - q_S - q_C = 0$ makes the objective function zero, the players can do better by setting prices such that $q_C - c_C \neq 0$ and $V + \delta_C - q_S - q_C \neq 0$. This implies $\lambda_2 = \lambda_3 = 0$.

B.1. Case I: $q_S \neq r$

$q_S \neq r$ implies $\lambda_1 = 0$. Solving (10) and (11) gives $(q_S, q_C) = \left(\frac{V+\lambda_C}{3}, \frac{V+\lambda_C}{3} + c_C\right)$ where $\lambda_C = \delta_C - c_C$. This is feasible only when $r \leq \frac{V+\lambda_C}{3}$. The expected payoffs for seller, buyer and consultant are same and is given by $\frac{V+\lambda_C}{3}$. The objective value for this case is $\left(\frac{V+\lambda_C}{3}\right)^3$.

B.2. Case II: $q_S = r$

From equation (11) we get $q_C = \frac{V+\lambda_C-r}{2} + c_C$ where $\lambda_C = \delta_C - c_C$. For a feasible solution, we need $q_C \geq c_C$ i.e. $V + \lambda_C \geq r$. The equilibrium expected payoffs of seller, buyer, and consultant are $\left(r, \frac{V+\lambda_C-r}{2}, \frac{V+\lambda_C-r}{2}\right)$. The objective value in this case is $r\left(\frac{V+\lambda_C-r}{2}\right)^2$.

Since both above cases are feasible in the range $r \leq \frac{V+\lambda_C}{3}$, we need to compare their objective values $\left(\frac{V+\lambda_C}{3}\right)^3$ and $r\left(\frac{V+\lambda_C-r}{2}\right)^2$ to find the global optimum. We compare the arithmetic mean and geometric mean of r, $\frac{1}{2}(V+\lambda_C-r)$, and $\frac{1}{2}(V+\lambda_C-r)$ i.e. $\frac{1}{3}\left(r + \frac{V+\lambda_C-r}{2} + \frac{V+\lambda_C-r}{2}\right) \geq \sqrt[3]{r\left(\frac{V+\lambda_C-r}{2}\right)\left(\frac{V+\lambda_C-r}{2}\right)}$ which implies $\left(\frac{V+\lambda_C}{3}\right)^3 \geq r\left(\frac{V+\lambda_C-r}{2}\right)^2$. Therefore, $\left(\frac{V+\lambda_C}{3}, \frac{V+\lambda_C}{3}, \frac{V+\lambda_C}{3}\right)$ is the global solution when $r \leq \frac{V+\lambda_C}{3}$ and $\left(r, \frac{V+\lambda_C-r}{2}, \frac{V+\lambda_C-r}{2}\right)$ is the global solution in the range $\frac{V+\lambda_C}{3} \leq r \leq V + \lambda_C$. For $r \geq V + \lambda_C$, none of the prices satisfy all the conditions, and therefore, there will be no agreement – the disagreement outcome $(r, 0, 0)$ will result. ■

C. Proof of Lemma 3

LEMMA 3. *When the buyer plans to hire a fixed price consultant, the equilibrium expected payoffs U_S^F and U_B^F of seller and buyer respectively from the negotiation are as follows where $\lambda_C^F = \delta_C^F - q_C^F \geq 0$*

$$(U_S^F, U_B^F) = \begin{cases} \left(\frac{1}{2}(V+\lambda_C^F), \frac{1}{2}(V+\lambda_C^F)\right), & \textit{if } r \leq \frac{1}{2}(V+\lambda_C^F) \\ \left(r, V+\lambda_C^F - r\right), & \textit{if } \frac{1}{2}(V+\lambda_C^F) \leq r \leq V+\lambda_C^F \\ (r, 0), & \textit{if } V+\lambda_C^F \leq r \end{cases}$$

Proof: The negotiation between seller and buyer with a fixed price consultant is given by 5a. The Lagrangian of the problem is $\mathcal{L}(q_S, \lambda_1, \lambda_2) = q_S(V' - q_S) - \lambda_1(q_S - r) - \lambda_2(V' - q_S)$ where $V' = V + \delta_C^F - q_C^F$. Rest of the proof follows similar to the seller and buyer sequential negotiation in appendix A. ■

D. Proof of Lemma 4

Lemma 4. *When the seller combines the data with data analytic services, the equilibrium expected payoffs U_S^C and U_B^C of the seller and buyer respectively form the negotiation are as follows where $\lambda_S = \delta_S - c_S \geq 0$ is the net contribution by the seller above data value.*

$$(U_S^C, U_B^C) = \begin{cases} \left(\frac{1}{2}(V+\lambda_S), \frac{1}{2}(V+\lambda_S), \right) & \textit{if } r \leq \frac{1}{2}(V+\lambda_S) \\ (r, V+\lambda_S - r,), & \textit{if } \frac{1}{2}(V+\lambda_S) \leq r \leq V+\lambda_S \\ (r, 0,), & \textit{if } V+\lambda_S \leq r \end{cases}$$

Proof: The negotiation between seller and buyer with seller as a consultant is given by 6a. We define $q_S' = q_S - c_C$ and $V' = V + \delta_S - c_S$. The problem 6a is now reduced to $\max_{r \leq q_S' \leq V'} q_S' \cdot (V' - q_S')$. This reduced problem is solved using the method described in appendix A for seller and buyer negotiation. ■