Monetizing Data Through B2B Negotiation: When is a Demonstration Appropriate?

Abstract

The explosive growth of eBusiness has allowed many companies to accumulate a repertoire of rich and unique datasets that can provide substantial value to other firms. Monetizing this data is a growing source of revenue to the data owners — one that can generate millions of dollars each year. The value of the data to a potential buyer is usually uncertain to both parties, and a mutually acceptable price is arrived at through negotiation. Evidence suggests buyers tend to underestimate value in the face of uncertainty. In this paper, we first analyze the problem of monetizing and selling proprietary data in the face of uncertainty as a Nash bargaining process, and find that the negotiation will divide the expected value of the data equally between the buyer and the seller. As the seller knows that the buyer is likely to be underestimating the value, they can choose to provide a demonstration to the buyer so they can arrive at a better estimate of dataset value. We extend our analysis by adapting the approach of Myerson (1984) to identify when such demonstrations are appropriate, and when they are not. Our results show that the extent of underestimation needs to be above a threshold for a demonstration to benefit the seller, and that the seller could be worse off providing a demonstration if the extent of underestimation is low.

1 Introduction

In a world where data is increasingly being collected at every imaginable opportunity, the real-time, proprietary data could be of immense value to many firms in making important business decisions — for example, consumer goods manufacturers have a lot to gain from point-of-sale data from retailers. Monga (2014) reports that Kroger, a well known supermarket operator, collects information on customer purchases at more than 2,600 stores, while tracking approximately 55 million loyalty-card members. This data is of substantial value to consumer goods manufacturers like Procter & Gamble and Nestlé, as it allows them to tailor their products and marketing to consumer preferences. Kroger’s revenues from the sale of
this data is estimated to be in the order of $100 million a year (Monga, 2014). This situation is not unique to the supermarket context. Travel technology companies like Sabre Holdings collect data on customer shopping activity, comprising of all the air travel related products considered for purchase by the travelers before finalizing on a specific purchase (Sabre Holdings, 2015). This data can provide airlines with considerable insights — an understanding of how well they are converting travel demand into bookings, for instance — that will enable them to make better decisions on new promotions, routes, schedules and capacity. Doshi (2015) estimates that “in the next three to five years, the first movers among airlines in data-driven personalization will quickly become the differentiated market leaders.”

The selling of such proprietary data is ubiquitous but does the seller and the buyer know how much is the data worth? This turns out to be a difficult question to answer, as data owners are seldom aware of the full potential of the data they hold. According to the Wall Street Journal “Data isn’t a physical asset like a factory or cash, and there aren’t any official guidelines for assessing its value” (Monga, 2014). However, this is a question that will be asked more often. — and the answer to this question could go a long way in providing data owners a solid basis for negotiation.

Unfortunately, while data owners know that their clients will be able to leverage their unique dataset to gain insights, they are often unclear on the price they can charge for it. Their knowledge of the client’s business needs is limited, and this limits the extent to which they can estimate the value of the data to the firms buying it. This raises a fundamental question — how should the data owners price their data? There is uncertainty on the data buyer’s side as well — while they know their business objectives, they have only limited knowledge about the data (for example, through a data dictionary provided by the data seller). Consequently, the buyer does not know how useful the data will be to them, which affects the price they are willing to pay for the data. In this B2B context, the price of the data is often arrived at through a negotiation process between the data seller and the client.
There are two main paradigms for selling data at present — selling by fixed price and selling through negotiations. The fixed price selling is adopted by data aggregators like aggdata.com who quotes a price and a time for data delivery to its clients. If the client agrees then an online contract is signed and aggdata.com start extracting information across the web based on their client’s request (https://www.aggdata.com/product-overview/custom-data-development). There also exist data selling platforms like thedataexchange.com who provide a platform through which sellers can find a market for their data. The price of the data can be fixed by the seller or can be arrived after a negotiation between the seller and the buyer (http://thedataexchange.com/selling-data/). The platform generates revenue by taking commission from each transaction. Apart from this there are data generators like Sabre Holdings and Kroger, who generate their own unique data by virtue of the services they provide. These firms can sell their data on a contractual basis after negotiations with the buyer.

This paper first identifies an incentive compatible pricing mechanism in a B2B data selling process through negotiation when the true value of the data is unknown to the players. We represent the negotiation as a Nash bargaining process (Nash, 1950), and show that the equilibrium price divides the expected value of the data equally between the seller and the buyer (Section 3.1). Prior evidence suggests that buyers often underestimate value in situations when the true value is unknown (Heiman and Muller, 1996). As a result, it might become necessary for the data seller to signal to the buyer that the data is worth more than their (under)estimated value. They can demonstrate the value of the data through dashboards that illustrate the benefits of the data (for example, in the form of analytical tools, summary statistics, relevant plots and decision criteria). In order to simplify our analysis, we assume that the demonstration by the data owner is effective, and that the buyer becomes fully aware of the value of the data as a result. On the other hand, the seller remains uncertain about the value of the data to the buyer, as no additional information
is conveyed by the buyer — that is, the result of the demonstration is only to clarify to
the buyer that they have underestimated the value of the data. Therefore, the decision
of whether or not a demo should be provided is an important one — this paper provides
insights into situations where a seller should (and should not) propose a free demo to the
buyer before the negotiation process. An incentive compatible decision mechanism adopted
from Myerson (1984) is proposed in the asymmetric information case when the buyer is fully
aware of the value and the seller is not (Section 3.4). Our analysis shows that extent of
underestimation affects the outcome. The extent of underestimation has to be substantial
enough for the seller to be better off by revealing the actual value to the buyer, and the
seller will be worse of if this is not the case (in which case, the seller should not be providing
a demo).

Section 2 provides a brief review of the literature on uncertainty in product valuation
and negotiation. Section 3 presents the two possibilities of the negotiation – Case 1, when
both the data seller and the buyer are unaware of the true value of the data, and Case 2,
when the buyer is aware of the true value but the seller is not. The data demonstration and
its effects on the negotiation is discussed in Section 4. Extension of the model are proposed
in Section 5 and Section 6 concludes.

2 Literature Review

The selling of products and services of uncertain value has been widely studied in both
business and economics. For example, Swinney (2011) discusses the effect of quick response
production practices when selling to a forward-looking consumer population that has un-
certain, heterogeneous valuations for a product; he modeled valuation as a private signal
distributed over the population where the quality of the signal determines the product val-
uation. Several papers have considered the situation where product value is revealed over
time. For instance, DeGraba (1995) looked at the monopolist behavior of artificially increasing demand by under-stocking, when customers who are initially uninformed about their valuation for a good realizes its true value over time. He modeled the valuation as a two point distribution of high and low over the population size. There is evidence that both the seller and the buyer are uncertain about product value when this value is realized at a point well after the product is purchased (Shugan and Xie, 2000). Our setting is similar to this context in that neither the data seller nor the data buyer are aware of the true value of the data at the time the contract is executed — the true value is revealed to the buyer at a later time, when they analyze the data to garner insights.

Another stream of relevant work is in the bargaining and negotiation domain. Nash (1950) provided the first formal investigation into bargaining when he looked at a static axiomatic theory of cooperative bargaining between two players with complete information. He showed that the equilibrium solution is that unique element from a set of alternatives that maximizes the product of utility payoffs to the players. This work precipitated a large volume of research on the bargaining process. Two among these deserve special mention — Harsanyi and Selten (1972) generalized the bargaining process for incomplete information where the types of the players are assumed to be private information, while Myerson (1984) extended it further by introducing a decision mechanism, whereby the players did not have to reveal their true types, but could agree on a mixed equilibrium over all the alternatives.

3 The Model

We consider a bargaining game between a data seller and a client/buyer. The seller has a unique dataset that is of value to the buyer. We consider two possibilities — Case 1, when the true value of the dataset is unknown to both the players and Case 2, when buyer knows the value that she would get from the data but the seller does not.
3.1 Case 1: True Value Unknown to Both Players

The true value of the dataset is unknown to the seller as they are not sure exactly what the buyer can use the data for. The buyer on the other hand is not sure of the extent to which the dataset will be useful in gleaning actionable insights. As a result, both the players are uncertain about the true value of the data. We assume however, that the distribution of dataset value is a common knowledge, and lies in the range \([V_L, V_H]\) \((V_L < V_H)\), with an expected value of \(V\). The buyer and seller enter a negotiation process to agree upon a mutually acceptable price for the dataset.

Let \(\mathcal{D}\) be the set of feasible alternatives that both the players can agree upon. Each element \(d \in \mathcal{D}\) denotes the actions taken by the players where seller quotes a price \(q_S(d)\) and buyer either accepts or rejects the quote. That is, the buyer has two alternatives for each price \(q_S\) — to accept the quote, or to reject it. We assume that the set \(\mathcal{D}\) is finite, compact and convex, and is known to both players. If the buyer accepts the quote \(q_S(d)\), the payoffs to the seller and buyer are \(u_S(d) = q_S(d)\) and \(u_B(d) = V - q_S(d)\) respectively. A rejection by the buyer naturally results on zero payoffs to both parties, i.e., \(u_S(d) = u_B(d) = 0\). The seller will not quote any price above \(V_H\), as they know that the buyer will reject that quote with certainty (since \(u_B(d)\) will be less than 0 if they accept). The cost of collecting the data is normalized to zero without loss of generality. Therefore, \(q_D\) will be selected from the range \([0, V_H]\). The set \(\mathcal{D}\) can be written as \(\mathcal{D} = \mathcal{D}_A \cup \mathcal{D}_R\), where \(\mathcal{D}_A\) is the set of quotes accepted by the buyer and \(\mathcal{D}_R\) is the set of quotes rejected by the buyer.

Rather than agreeing upon a single alternative, suppose the players agree on a mixed strategy probability distribution over the set of alternatives. That is, the players decide on the distribution \(\mu(d)\) which is the probability of selecting the alternative \(d\) such that \(\sum_{d \in \mathcal{D}} \mu(d) = 1\) and \(\mu(d) \geq 0\) for all \(d \in \mathcal{D}\). The expected payoff to player \(i \in \{S, B\}\) in a mixed strategy is given by \(U_i = \sum_{d \in \mathcal{D}_A} \mu(d)u_i(d)\). Since \(u_S(d) \geq 0\) for each \(d\) hence \(U_S \geq 0\). We also need the individual rationality constraint to be satisfied, i.e. \(U_i \geq 0\). The outcome
of the negotiation is the mixed strategy equilibrium \( \mu^* \) that maximizes the Nash product (i.e. the product of expected payoffs, \( U_S \cdot U_B \)) (Nash, 1950). The equilibrium strategy \( \mu^* \) will be selected from a set of feasible strategies \( \psi \).

\[
(NN) \quad \max_{\mu \in \psi} \left( \sum_{d \in \mathcal{D}_A} \mu(d)u_S(d) \right) \left( \sum_{d \in \mathcal{D}_A} \mu(d)u_B(d) \right) \tag{1}
\]

s.t.
\[
\sum_{d \in \mathcal{D}_A} \mu(d)u_S(d) \geq 0; \quad \sum_{d \in \mathcal{D}_A} \mu(d)u_B(d) \geq 0 \tag{2}
\]

\[
\sum_{d \in \mathcal{D}} \mu(d) = 1; \quad \mu(d) \geq 0; \quad \forall d \in \mathcal{D} \tag{3}
\]

Proposition 1 states that the solution to \((NN)\) will divide the expected value equally between the two players.

**Proposition 1.** The negotiation between seller and buyer will end up giving equal expected payoffs, \( \frac{V}{2} \) to the players when the distribution of the value is a common knowledge with expected value of \( V \).

**Proof.** Since \( u_B(d) = V - u_S(d) \) for all \( d \in \mathcal{D}_A \) therefore we can write \( U_B \) as,

\[
U_B = \sum_{d \in \mathcal{D}_A} \mu(d)(V - u_S(d))
\]

\[
= \delta V - U_S \tag{4}
\]

Where \( \delta = \sum_{d \in \mathcal{D}_A} \mu(d) \). Since \( U_S \geq 0 \) hence equation (4) gives an upper bound of \( \delta V \) for \( U_B \). Therefore, the maximization problem will make the variable \( \delta = 1 \) which implies that all the rejection agreements \( d \in \mathcal{D}_R \) will have a zero probability of being selected by the mechanism \( \mu \). The negotiation problem \((NN)\) can now be transformed in terms of \( U_S \) and \( U_B \) as,

\[
\max_{U_S,U_B} U_S, U_B \tag{5}
\]

s.t.
\[
U_S + U_B = V \tag{6}
\]

\[
U_S, U_B \geq 0 \tag{7}
\]

Let \( f(U_S) = U_S(V - U_S) \). The first order necessary conditions gives \( \frac{\partial f}{\partial U_S} = V - 2U_S = 0 \). Hence the critical point is \( \frac{V}{2} \). Furthermore, \( \frac{\partial^2 f}{\partial U_S^2} = -2 < 0 \) hence the second order sufficiency condition is satisfied. Therefore, \( f(U_S) \) attains its maximum at \( \frac{V}{2} \). This implies that the unique solution to the objective function \((NN)\) is \( \left( \frac{V}{2}, \frac{V}{2} \right) \) which is also the equilibrium payoffs of the negotiation. We can get the equilibrium mechanism \( \mu^* \) by
solving $U_i = \frac{V}{2} = \sum_{d \in D_A} \mu(d) u_i(d)$ for $i \in \{S, B\}$ along with the feasibility constraints (3). The mechanism $\mu^*$ will not be unique if the number of alternatives is more than the number of equations.

The Proposition 1 is to be expected, as both players are equally unaware of the true value and hence, are in a similar situation. Section 3.2 illustrates this proposition through an example.

### 3.2 Numerical Example 1

Let the distribution of value be $V_L = 2$ and $V_H = 6$ with probabilities $p_L = 0.2$ and $p_H = 0.8$ respectively; the expected value $V = p_L V_L + p_H V_H = 5.2$. Both players are aware of this distribution. For simplicity we assume that the seller can quote only integer prices and her reservation price is zero. The seller will select prices from the range $[0, 6]$, i.e $q_S$ can take 7 value $\{0, 1, 2, 3, 4, 5, 6\}$. For each of these prices the buyer can take an action of either accepting or rejecting the quote. Therefore, the set of alternatives will consists of the 14 elements given in Table 1. The elements of $\mathcal{D}$ can be divided into two mutually exclusive and collectively exhaustive sets $\mathcal{D}_A$ and $\mathcal{D}_R$ where $\mathcal{D}_A$ contains the alternatives accepted by the buyer and $\mathcal{D}_R$ contains the alternatives rejected by the buyer. Each element $d \in \mathcal{D}$ is represented by the pair $(q_S(d), A)$ or $(q_S(d), R)$ where $A$ and $R$ denotes B’s action of acceptance or rejection respectively.

<table>
<thead>
<tr>
<th>Set of Alternatives $\mathcal{D}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,R)</td>
</tr>
<tr>
<td>$u_S$</td>
</tr>
<tr>
<td>$u_B$</td>
</tr>
<tr>
<td>$\mu^*(d)$</td>
</tr>
</tbody>
</table>

Table 1: Payoff table for Case 1

The players will get a payoff of $u_S(d) = q_S(d)$ and $u_B(d) = V - q_S(d)$ if the alternative $d \in \mathcal{D}_A$ while the payoffs will be 0 if $d \in \mathcal{D}_R$. The decision variable is the mixed probability $\mu(d)$ which will be used to compute the expected payoff of each player. A vector of the feasible equilibrium solution $\mu^*$ is provided in the last row of Table 1. The expected payoffs
are \( U_S = \sum_{d \in D_A} \mu(d)u_S(d) = 0.53 \times 0 + 0.031 \times 4 + 0.157 \times 5 + 0.282 \times 6 = 2.6 \) and \( U_B = \sum_{d \in D_A} \mu(d)u_B(d) = 0.53 \times 5.2 + 0.031 \times 1.2 + 0.157 \times 0.2 + 0.282 \times (-0.8) = 2.6 \). Hence, in the Nash equilibrium the expected value of 5.2 is equally divided between the players, as suggested by Proposition 1.

### 3.3 Case 2: Buyer Knows the True Value

In Case 1, we found that the players divide the expected payoff equally when they only know the distribution of the value. In this section we will analyze another possibility where the buyer becomes aware of the value they will get from the data after looking at the data dictionary while the seller still remains unaware of the true value as they have a limited knowledge about the buyers’ requirements. Let the distribution of the value be \( V_L \) with probability \( p_L \) and \( V_H \) with probability \( p_H \) \((= 1 - p_L)\) respectively, with expected value of \( V = p_LV_L + p_HV_H \). We assume that the data description provided by the seller helps the buyer to recognize whether they will realize a value of \( V_H \) or \( V_L \). The buyer is of type ‘High’ \((H)\) if they realize a value of \( V_H \) and of type ‘Low’ \((L)\) if they realize a value of \( V_L \) while the seller still assumes it to follow the two point distribution of \( V_L \) and \( V_H \) with probability \( p_L \) and \( p_H \) respectively. Let \( T_B = \{t_B\} \) be the set of buyer types, where \( t_B = L, H \). Given this information, the players will start negotiating; this calls for a bargaining game under asymmetric information. Before analyzing this game we need to modify the mixed strategy decision variable stated in Section 3.1 by a decision mechanism which will incorporate the buyer types.

### 3.4 The Decision Mechanism

Following Myerson (1984), we assume that the players agree on a decision rule or mechanism \( \mu(d|t_B) \) which is a function on the domain \( D \times T_B \) defined by
\[
\sum_{d \in \mathcal{D}} \mu(d|t_B) = 1; \quad \forall t_B \in \mathcal{T}_B \quad (8)
\]
\[
\mu(d|t_B) \geq 0; \quad \forall d \in \mathcal{D}, \quad t_B \in \mathcal{T}_B \quad (9)
\]

The randomized strategy \(\mu(d|t_B)\) is the probability of selecting alternative \(d\) in the mechanism \(\mu\), given the buyer type \(t_B\). In the bargaining game both players will agree on a mechanism \(\mu(d|t_B)\) to be implement when type \(t_B\) is revealed by the buyer after the negotiation. This encourages the buyer to agree on mechanism \(\mu\) without actually revealing her true type. The payoffs to the seller and the buyer in each alternative \(d\) is given by \(u_B(d, t_B) = V_{t_B} - q_S(d)\), \((t_B = L, H)\), if \(d \in \mathcal{D}_A\) and 0 otherwise. Also, \(u_S(d) = q_S(d)\) if \(d \in \mathcal{D}_A\) and 0 otherwise, irrespective of buyer’s type. Let \(U_B(\mu, s_B|t_B)\) be the conditional expected payoff for the buyer of type \(t_B\) while pretending to be of type \(s_B\) \((\forall t_B, s_B \in \mathcal{T}_B)\), when mechanism \(\mu\) is implemented. Also, let \(U_S(\mu)\) be the expected payoff for the seller when the mechanism \(\mu\) is implemented. These expected payoffs can be calculated as follows
\[
U_B(\mu, s_B|t_B) = \sum_{d \in \mathcal{D}_A} \mu(d|s_B)u_B(d, t_B); \quad \forall t_B, s_B \in \mathcal{T}_B \quad (10)
\]
\[
U_S(\mu) = \sum_{d \in \mathcal{D}_A} (p_L\mu(d|L) + p_H\mu(d|H))u_S(d); \quad (11)
\]

A mechanism is said to be feasible if it satisfies conditions (8) and (9). It also needs to be Bayesian Incentive Compatible (BIC) and Individually Rational (IR) i.e., \(\mu\) must also satisfy the following conditions
\[
U_B(\mu, L|L) \geq U_B(\mu, H|L); \quad U_B(\mu, H|H) \geq U_B(\mu, L|H) \quad (12)
\]
\[
U_B(\mu, t_B|t_B) \geq 0; \quad U_S(\mu) \geq 0; \quad \forall t_B \in \mathcal{T}_B \quad (13)
\]

Let \(\psi\) be the set of feasible mechanisms for the bargaining game. According to Harsanyi and Selten (1972) and Myerson (1984), the solution to this Bayesian bargaining game is the mechanism \(\mu \in \psi\) that maximizes the generalized Nash product, subject to the constraints discussed earlier.
\[(NY) \quad \max_{\mu \in \Psi} \quad U_S(\mu) (U_B(\mu, L|L))^{p_L} (U_B(\mu, H|H))^{p_H} \tag{14}\]

\[\text{s.t.} \quad U_B(\mu, L|L) \geq U_B(\mu, H|L); \quad U_B(\mu, H|H) \geq U_B(\mu, L|H) \tag{15}\]

\[U_S(\mu) \geq 0 \tag{16}\]

\[U_B(\mu, L|L) \geq 0; \quad U_B(\mu, H|H) \geq 0 \tag{17}\]

\[\sum_{d \in D} \mu(d|L) = 1; \quad \sum_{d \in D} \mu(d|H) = 1 \tag{18}\]

\[\mu(d|L) \geq 0; \quad \mu(d|H) \geq 0; \quad \forall d \in D \tag{19}\]

Steps similar to the proof of Proposition 1 is used to convert \(U_S(\mu), U_B(\mu, L|L)\) and \(U_B(\mu, H|H)\) in terms of \(S_L(\mu) = \sum_{d \in D_A} \mu(d|L)u_S(d)\) and \(S_H(\mu) = \sum_{d \in D_A} \mu(d|H)u_S(d)\) as,

\[U_S(\mu) = p_L S_L + p_H S_H \tag{20}\]

\[U_B(\mu, L|L) = \frac{1}{V_H} (V_L(V_H - S_H) - S_L(V_H - V_L)) \tag{21}\]

\[U_B(\mu, H|H) = V_H - S_H \tag{22}\]

After some algebra, we can write \((NY)\) in a following simple form.

\[\max_{S_L, S_H} \quad (p_L S_L + p_H S_H) (V_L(V_H - S_H) - S_L(V_H - V_L))^{p_L} (V_H - S_H)^{p_H} \tag{23}\]

\[\text{s.t.} \quad S_L \geq 0; \quad S_H \geq S_L \tag{24}\]

The above transformed form is solved to get \(S_L(\mu^*)\) and \(S_H(\mu^*)\) which in turn gives \(U_S(\mu^*), U_B(\mu^*, L|L)\) and \(U_B(\mu^*, H|H)\) from equations (20), (21) and (22). We can now use the equations of \(S_L(\mu^*), S_H(\mu^*)\) and (10) along with the feasibility equations (8) and (9) to get the mechanism \(\mu^*\). The \(\mu^*\)-vector will not be unique if the number of alternatives is more than the number of equations. Proposition 2 gives the expected payoffs to the players as an outcome of the negotiation. (Proofs of propositions 2 and 3 are excluded for the sake of brevity).

**Proposition 2.** When the buyer knows the true value of the data while the seller assumes it to be distributed as \(V_L\) with probability \(p_L\) and \(V_H\) with probability \(p_H\), the outcome of the
negotiation will give following expected payoffs to the players,

\[(U_S^*, U_B^*) = \begin{cases} 
(pHV_H, \frac{V_L}{2}), & \text{if } \frac{V_L}{V_H} \leq p_H \leq 1 \\
(pHV_H, \frac{(V-V_L)(p_HV_H-p_LV_L)}{2(p_HV_H-p_LV_L)}), & \text{if } \frac{2V_L}{2V_H-V_L} < p_H < \frac{V_L}{V_H} \\
(D_s, V-D_s), & \text{if } 0 \leq p_H \leq \frac{2V_L}{2V_H-V_L} \end{cases} \]

where, \(D_s = \frac{1}{4} \left(2(V_L + V_H) - V - \sqrt{(2(V_L + V_H) - V)^2 - 8V_LV_H} \right)\) and \(V = p_LV_L + p_HV_H\)

In the above proposition, \(U_S^* = U_S(\mu^*)\) and \(U_B^* = p_LU_B(\mu^*, L|L) + p_HU_B(\mu^*, H|H)\). What does this tell us about the impact on the players? That is, who is better off in \((NY)\) compared to \((NN)\)? To check this, we need to compare the expected payoffs in Propositions 1 and 2. It can be easily verified that the seller is worse off and the buyer is at least as well off as before in the \((NY)\) case. We will look at an example to better understand Proposition 2.

### 3.5 Numerical Example 2

As in Example 1 from section 3.2, let the distribution of the value be \(V_L = 2\) and \(V_H = 6\) with probabilities \(p_L = 0.2\) and \(p_H = 0.8\) respectively, with expected value \(V = 5.2\). We assume that the buyer is aware of the value while the seller still assumes it to follow the original two point distribution. The buyer is of type \(H\) if the value is 6 and of type \(L\) if the value is 2. The set of alternatives will consist of the same 14 elements as before given in Table 2.

<table>
<thead>
<tr>
<th>Set of Alternatives (D)</th>
<th>(0,R)</th>
<th>(0,A)</th>
<th>(1,R)</th>
<th>(1,A)</th>
<th>(2,R)</th>
<th>(2,A)</th>
<th>(3,R)</th>
<th>(3,A)</th>
<th>(4,R)</th>
<th>(4,A)</th>
<th>(5,R)</th>
<th>(5,A)</th>
<th>(6,R)</th>
<th>(6,A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(u_S)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>(u_B(L))</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>-2</td>
<td>0</td>
<td>-3</td>
<td>0</td>
<td>-4</td>
</tr>
<tr>
<td>(u_B(H))</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(\mu^*(d</td>
<td>L))</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(\mu^*(d</td>
<td>H))</td>
<td>0</td>
<td>0.372</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.125</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0590</td>
<td>0</td>
<td>0.155</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2: Payoff table for Case 2

Table 2 also shows the payoffs received by the players in each alternative. When \(d \in D_A\) the seller will get \(u_S(d) = q_S(d)\) and the buyer will get \(u_B(d|t_B) = V_{t_B} - q_S(d)\) depending on whether the buyer is a high type \((t_B = H)\) or a low type \((t_B = L)\). Payoffs in all
other alternatives will be zero. The decision variable is the mechanism $\mu(d|t_B)$ for type $t_B = L, H$. Since, $p_H(=0.8)$ lies between $\frac{V_L}{V_H-V_L}(=0.5)$ and 1, hence, according to the proof of Proposition 2, $S_L(\mu^*) = 0$ and $S_H(\mu^*) = \frac{V_H}{2}$. This gives $U^*_S = \frac{p_U V_H}{2} = 2.4$ and $U^*_B = \frac{V}{2} = 2.6$. The feasible equilibrium mechanism $\mu^*$ can be obtained by solving the equations $S_L(\mu^*)$, $S_H(\mu^*)$ and (10) along with the feasibility equations (8) and (9). One such vector of the feasible equilibrium solution $\mu^*$ is provided in Table 2. This example illustrates that the seller is worse off in Case 2. As expected, the buyer cannot be worse off after being made aware of the value.

4 Providing a Demonstration

In Case 2, we found that the seller is worse off when the buyer has the full knowledge of the value of the data. As mentioned earlier, buyers are known to underestimate value when the true value is not known (Heiman and Muller, 1996). As underestimation of value on the part of the buyer reduces the price they would be willing to pay, the seller might wish to provide some evidence of the benefits of the data to the buyer before the negotiation process. A common way to do this is through a demonstration that clarifies the true value of the data. Typically, these demonstrations involve showcasing a dashboard that contains some results of mining the data. Such a demonstration can reduce the uncertainty about the value of the data, and provide the buyer with a better idea of its potential. After the demonstration, the buyer will be in a better position to evaluate the potential benefit of the data in terms of insights they can gain from it. For simplicity, we assume a demonstration will reveal the exact value of the data to the buyer. While the seller’s knowledge has not changed — the seller is still uncertain about the exact value of the data to the buyer — they know now that the buyer is no longer underestimating the value of the dataset. We have also assumed that preparing and showcasing of the demo is a costless affair for the seller.
In the following Sections 4.1, we will analyze whether proposing such a free demonstration before the negotiation process is beneficial to the seller.

4.1 Demo Reduces Underestimation

Here we consider the situation when the buyer realizes that the value could be as high as \( V_H^+ > V_H \) after attending the demonstration. This is to say that the demo not only helps the buyer to realize their type, but also reduces the extent by which they underestimate the value. The seller also realizes that the demo has improved the upper bound of the distribution to \( V_H^+ \) and updates the distribution to \( V_L \) and \( V_H^+ \) with probabilities \( p_L \) and \( p_H \) respectively. Now the seller is interested in what \( V_H^+ \) needs to be before a demonstration is warranted, that is, what improvement in the upper bound is needed so that they get an expected payoff greater than \( \frac{V}{2} \), the expected payoff when no demo was provided? This question is answered by Proposition 3.

**Proposition 3.** If the demonstration increases the upper bound of the valuation from \( V_H \) to \( V_H^+ \) while keeping the lower bound and the probabilities same as before, then showing a demo will increase the expected payoff of the seller when following criteria are met:

\[
V_H^+ > \begin{cases} \frac{V}{p_H}, & \text{if } \frac{2V_L}{2V_H-V_L} < p_H \leq 1 \\ V^B = V \left( \frac{2V_L-2V_H-p_L(V_H-V_L)}{2V_H-V_L} \right), & \text{if } \frac{2V_H-3V_L}{V_H-V_L} < p_H \leq \frac{2V_L}{2V_H-V_L} \text{ and } (1 - 1/\sqrt{3})V_H < V_L < V_H \end{cases}
\]

where, \( V = p_L V_L + p_H V_H \).

Proposition 3 is best understood through Figure 1, which plots the conditions relating to when the seller should propose a demo. Figure 1 implies that if the probability of the value being high is small, i.e., \( p_H < \min \left( \frac{2V_L}{2V_H-V_L}, \frac{2V_H-3V_L}{V_H-V_L} \right) \), and \( V_H \) is sufficiently high compared to \( V_L \) (i.e., \( \frac{V}{V_H} < \frac{2}{3} \)) then the seller is always worse off by showing the demo. In this situation there is a high probability that the value is low, and revealing that to the buyer through a demonstration will be detrimental for the seller. When \( V_H \) is close enough to \( V_L \) (i.e., \( \frac{V}{V_H} > \frac{2}{3} \)) and the demo reduces the extent of underestimation (\( V_H^+ > V^H \)) then providing a
demonstration before the negotiation process becomes a viable option for the seller.

![Figure 1: Criteria for Demo Proposal]

**4.2 Numerical Example 3**

In this example we apply Proposition 3 to the Example 2 from Section 3.5 to find a workable demo that would give the seller a higher expected payoff than \( \frac{V}{2}(=2.6) \). Since \( p_H(=0.8) \) lies between \( \frac{2V_L}{2V_H-V_L}(=0.4) \) and 1, Proposition 3 implies that the demo will need to indicate the upper bound of the value to be greater than \( \frac{V}{p_H}=6.5 \) to make it a viable option for the seller. Let us assume that the demonstration improves \( V_H \) from 6 to 7 while keeping other parameters the same as before. After the demonstration the buyer is of type \( H \) with the value of 7 or of type \( L \) with the value of 2. The set of alternatives will now consists of the 16 elements given in Table 3. The expected payoffs after solving the problem (NY) using Proposition 2 are \( U^*_S=2.8 \) and \( U^*_B = p_L U_B(\mu^*, L|L) + p_H U_B(\mu^*, H|H) = 3 \). It is evident

<table>
<thead>
<tr>
<th>( u_S )</th>
<th>(0,R)</th>
<th>(0,A)</th>
<th>(1,R)</th>
<th>(1,A)</th>
<th>(2,R)</th>
<th>(2,A)</th>
<th>(3,R)</th>
<th>(3,A)</th>
<th>(4,R)</th>
<th>(4,A)</th>
<th>(5,R)</th>
<th>(5,A)</th>
<th>(6,R)</th>
<th>(6,A)</th>
<th>(7,R)</th>
<th>(7,A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( u_B(L) )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>( u_B(H) )</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>-2</td>
<td>0</td>
<td>-3</td>
<td>0</td>
<td>-4</td>
<td>0</td>
<td>-5</td>
<td></td>
</tr>
<tr>
<td>( \mu^*(d,L) )</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \mu^*(d,H) )</td>
<td>0</td>
<td>0.397</td>
<td>0</td>
<td>0.059</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.003</td>
<td>0</td>
<td>0.090</td>
<td>0</td>
<td>0.177</td>
<td>0</td>
<td>0.273</td>
</tr>
</tbody>
</table>

**Table 3: Payoff Table After Demonstration**

Proposition 2 are \( U^*_S=2.8 \) and \( U^*_B = p_L U_B(\mu^*, L|L) + p_H U_B(\mu^*, H|H) = 3 \). It is evident
that showing the free demo has increased the seller’s expected payoff from 2.6 to 2.8. Also, the buyer is always better off after attending a data demonstration — in this example, the buyer’s expected payoff has increased from 2.6 to 3.

5 Future Extension: Consultant as a Gatekeeper

For a variety of reasons (e.g., inadequate in-house capability, or to save cost), buyers often hire consultants to analyze the data for them rather than doing it in-house. The consultant works with the buyer as a gatekeeper. In this case, the buyer will need to pay the seller for the data and the consultant for their services. These prices are the outcome of a tripartite negotiation among the data seller, the buyer and the consultant. Proposition 1 can easily be extended to show that the negotiation will divide the expected value equally among all three players when none of them are aware of the true value of the data. The opportunity of providing a demonstration in this situation raises some interesting questions — should the seller propose a free data demonstration to the buyer knowing that the buyer may also ask the consultant to attend it? If the seller proposes a demonstration, should the buyer ask the consultant to attend it, knowing that a demo will make both the buyer and the consultant aware of the value of the data? Is it beneficial for the seller if the consultant also attends the demonstration? We intend to address these questions in a future extension to the models presented in this paper.

6 Conclusion

We consider a data selling negotiation process in a B2B context. We initially consider a situation where both the seller and buyer are unaware of the true value of the data and solve the problem by maximizing the Nash product of the expected payoffs of the players (Nash,
We consider a mixed strategy equilibrium over a set of feasible alternatives and find that the negotiation process will divide the expected value between the players equally. We then consider a case where the seller can increase her expected profit by providing a free demonstration before the negotiation process when the buyer is underestimating dataset value. The demonstration is found to be beneficial to the seller if it increases the upper bound of the value when the probability of getting high value is not too low. The buyer will never be worse off by attending the demonstration. We plan to extend this work by considering a context where the buyer hires a consultant to analyze the data, which results in a tripartite negotiation process.

References


