9-4-2020

Pricing and Quality Provision in a Supply Relationship: A Model of Efficient Relational Contracts

Cristina Nistor
Chapman University, nistor@chapman.edu

Matthew Selove
Chapman University, selove@chapman.edu

Follow this and additional works at: https://digitalcommons.chapman.edu/business_articles

Recommended Citation

This Article is brought to you for free and open access by the Business at Chapman University Digital Commons. It has been accepted for inclusion in Business Faculty Articles and Research by an authorized administrator of Chapman University Digital Commons. For more information, please contact laughtin@chapman.edu.
Pricing and Quality Provision in a Supply Relationship: A Model of Efficient Relational Contracts

Cristina Nistor, Matthew Selove

1. Introduction

Many business transactions involve complex agreements that would be difficult to enforce with a formal contract. In this context, relationships become crucial: the value of the future transactions gives each party in a relationship an incentive to perform at a level beyond that which would be possible to enforce in a formal contract.

For example, suppliers of high-end restaurants often customize food ingredients to meet the particular needs of each restaurant they serve. Writing down precise details such as the storage temperature or the cut of each ingredient would require a long, complex contract that a court might find practically difficult to enforce. However, if a supplier has an ongoing relationship with a restaurant, the two firms can establish an informal agreement that the supplier will make reasonable effort to satisfy the restaurant’s requests, and in return, the restaurant will pay the supplier a price premium over market rates for its ingredients. As long as the value of the relationship to each party exceeds the cost of the additional effort or payments they make, the relationship is sustainable.

This paper formulates a model of payment schemes that will sustain a relationship between a supplier and its downstream partner at efficient quality levels. Our model is a repeated game in which a supplier sells multiple goods to a downstream firm. In each period, the supplier can add a dimension of quality, such as product customization, to one of the goods. However, firms cannot implement formal contracts in which payment is contingent on product quality, for example, because it would be too difficult for firms to specify and for courts to enforce such contracts (Iyer and Villas-Boas 2003). We show how channel members can use relational contracts (Baker et al. 2002) to sustain these uncontractible services. The supplier’s desire to sustain the relationship provides an incentive to exert effort. Furthermore, the downstream player pays higher wholesale prices than offered on the outside market in order to reward the supplier’s past effort and sustain the relationship.

Our model focuses on the empirically common case of industries with unit prices that remain stable over time. Previous research has documented several reasons for stable prices, including price adjustment costs that are large enough to discourage frequent price changes (Zbaracki et al. 2004, Levy et al. 2010, Maciejovsky and Wernerfelt 2011) and fairness concerns that also compel the firm not to vary prices over time (Anderson and Simester 2010, Selove 2019).
We show how firms in a supply relationship can set prices for multiple goods to achieve pricing flexibility without needing to adjust prices as demand fluctuates. We then extend our model (in Section 3.4) to allow for costly adjustments to the total payment in a given period; in this case, firms would like to set per-unit prices that generate optimal payments during the most likely states of demand and incur adjustment costs during less likely states.

If some goods are customizable and some are not, we might expect that the most efficient relationship (i.e., the relationship that generates the greatest possible surplus) would set a price premium only on the customizable goods to reflect the greater value of effort in times of high demand for these goods. However, if demand for the customizable goods is very volatile, this creates a problem. When demand is very high, the premium payment the retailer is required to make may be so high that it exceeds the value of the effort that the supplier is willing to exert given the (future) value of the relationship.

Firms face a dilemma between wanting to provide higher total premium payments to reflect the greater value of effort in times of high demand and, at the same time, not wanting to require unsustainable premium payments or unsustainable effort levels during those periods. We show that in some cases, setting price premiums on both a customizable good and a non-customizable good can help resolve this dilemma. If demand for different types of goods is correlated, but demand for noncustomizable goods is less volatile, then a price premium on noncustomizable goods can serve the dual role of providing the supplier with higher rewards in times of high demand while still limiting the spikes in these payments to a range that is acceptable to the retailer.

We interviewed a sales manager for a distributor that supplies fish, rice, and other inputs to sushi restaurants. The manager told us that some restaurants require frequent customization of fish. He considers them good clients who will pay more overall as they give him more future business and buy other products from his company.

The manager provided us with a list of restaurants’ requests, for example, for fish to be “clean and white, please” or for a specific size and cut of fish (“half loin, please, front part” or for “15 lb back loin, head part, please !!!!!!”) or even for how frozen the items should be (“Be super frozen tuna saku—1 bag BIGG” or “2 lb not super frozen”). See Table 1 for a list of all customizable products sold by the distributor during an 11-month period. Instead of using formal contracts to ensure these requests are satisfied, the manager said he was responsible for maintaining informal personal relationships with restaurant owners, who are willing to pay premium prices if the supplier consistently provides customized ingredients.

The sushi supply manager we interviewed said that restaurants that require frequent customization of fish typically buy multiple different products from the supplier. Some of these restaurants pay a price premium on customized products such as salmon, some pay a premium on both customized and noncustomized products, and some actually compensate the supplier mostly with a price premium on noncustomized products such as a rice or seaweed. For an example of two products that exhibit this pattern, please see Online Appendix A. This pricing arrangement is consistent with our model, because products such as fresh fish have volatile demands, so any payment scheme linked only to the actual product being customized might not be able to sustain the relational contract that provides the customization effort. By contrast, dry products such as rice or seaweed are relatively stable goods that are perfect for serving as a basis for the premium payment to sustain the relational contract. Thus, our model provides one possible explanation for cross-subsidized pricing, in which some restaurants that require customization of their fish might pay a premium price for rice.

Arrow Electronics is another example of a company that uses this type of pricing structure to sustain a relationship with its customers. The company is a distributor of two types of microchips: standard commodity chips and value-added chips the company customizes for each client. The latter type of products requires effort investment by Arrow. The company forms relationships with clients who need customization, and these clients do not pay a price premium for value-added products that are customized but instead pay higher margins on the standard chips. If these customers go to the outside market for the commodity products, their relationship with Arrow typically ends (Narayandas 1998).

In our model, prices and effort levels in an efficient relationship depend on the amount of surplus generated by marginal effort (the surplus generated is the buyer’s per-unit benefit of effort minus the supplier’s per-unit cost of effort). If effort generates a moderate amount of surplus (surplus is not too low or too high), firms can sustain a relationship, but they cannot sustain maximum effort during large demand spikes. Therefore, a price premium only on the customizable good is not sustainable because it would lead to payment spikes that exceed the value of effort during high demand.

In such cases, we show that the most efficient relationship sets a price premium on multiple goods or, in some cases, only on a noncustomizable good with relatively stable demand to achieve the precise payment level during each demand state that equals the value of sustainable effort during that demand state. As the volatility of the customizable good increases, firms need to shift more of the price premium from this volatile good to a more stable good in order to maintain a given difference in the total premium payment between high- and low-demand states.
We also present model extensions in which it is not possible to choose prices that always generate the desired payment levels, and instead firms incur payment adjustment costs during certain periods (Section 3.4) or split their order quantity between their main supplier and the outside market to limit payment spikes during high demand (Section 3.5).

Our model implies that contracts become more efficient when firms have more instruments available to match their desired payment levels in each demand state. Therefore, firms would generally like to use all available pricing instruments if there is no additional transaction cost to doing so. For example, if firms can include a fixed-payment component, similar to two-part tariffs, then doing so makes their contract more efficient. However, we show (Section 3.6) that two-part tariffs do not eliminate the benefit of also setting a per-unit price premium on multiple goods. In particular, if there are more than two possible quantity levels for the customizable good, then a two-part tariff does not allow firms to set payment levels that increase linearly up to a certain quantity and then become less sensitive to quantity above that level. In such cases, firms can use per-unit price premiums on multiple goods to provide additional degrees of freedom in determining the total payment in each period.

An important implication of our model is that trading multiple goods can increase equilibrium profits for both firms by allowing more flexible payment arrangements than would be possible if they traded only the customizable good. Therefore, a supplier that provides customized goods may want to offer a full product line, including goods that do not require customization. Buyers that require customization then have an incentive to purchase all their goods from the same supplier, and both firms can benefit from this arrangement.

Section 2 reviews related literature. Section 3 presents the theoretical model. Section 4 concludes with implications for channel management and directions for future research. Online Appendix A contains an

### Table 1. Customizable Items

<table>
<thead>
<tr>
<th>Product name</th>
<th>Percent of total orders</th>
<th>Percent of customized orders</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frozen Scottish Whole Salmon</td>
<td>9.09</td>
<td>27.64</td>
</tr>
<tr>
<td>Y/F Tuna Loin</td>
<td>3.68</td>
<td>11.20</td>
</tr>
<tr>
<td>Fluke (Hirame)</td>
<td>2.64</td>
<td>8.03</td>
</tr>
<tr>
<td>Farm Rock</td>
<td>2.53</td>
<td>7.69</td>
</tr>
<tr>
<td>Frozen Escolar Block</td>
<td>2.20</td>
<td>6.70</td>
</tr>
<tr>
<td>California Uni</td>
<td>1.83</td>
<td>5.56</td>
</tr>
<tr>
<td>Frozen Smoked Salmon Chunk</td>
<td>1.81</td>
<td>5.51</td>
</tr>
<tr>
<td>Frozen Hamachi Fillet</td>
<td>1.66</td>
<td>5.06</td>
</tr>
<tr>
<td>(Frozen) Scottish Salmon Fillet for Sushi</td>
<td>0.85</td>
<td>2.60</td>
</tr>
<tr>
<td>Big Eye Tuna Loin</td>
<td>0.78</td>
<td>2.38</td>
</tr>
<tr>
<td>Frozen Hamachi Loin Farm Japan</td>
<td>0.55</td>
<td>1.69</td>
</tr>
<tr>
<td>Asi Beff Gyoza</td>
<td>0.51</td>
<td>1.55</td>
</tr>
<tr>
<td>Mushidako Octopus</td>
<td>0.49</td>
<td>1.48</td>
</tr>
<tr>
<td>Fresh Hamachi Fillet (Japan)</td>
<td>0.43</td>
<td>1.30</td>
</tr>
<tr>
<td>Aji</td>
<td>0.43</td>
<td>1.29</td>
</tr>
<tr>
<td>Frozen Albacore Tuna Loin</td>
<td>0.42</td>
<td>1.28</td>
</tr>
<tr>
<td>Frozen Escolar Block</td>
<td>0.41</td>
<td>1.25</td>
</tr>
<tr>
<td>Live Mirugai (Geoduck)</td>
<td>0.41</td>
<td>1.23</td>
</tr>
<tr>
<td>Madai (Japan)</td>
<td>0.29</td>
<td>0.87</td>
</tr>
<tr>
<td>Fresh Kanpachi Fillet</td>
<td>0.24</td>
<td>0.73</td>
</tr>
<tr>
<td>Unagi</td>
<td>0.24</td>
<td>0.72</td>
</tr>
<tr>
<td>Spanish Mackeral (USA)</td>
<td>0.23</td>
<td>0.69</td>
</tr>
<tr>
<td>OO-Toro Southern Blufin</td>
<td>0.22</td>
<td>0.68</td>
</tr>
<tr>
<td>Atlantic Whole Salmon (Farm Raised)</td>
<td>0.21</td>
<td>0.62</td>
</tr>
<tr>
<td>Tuna Ground</td>
<td>0.20</td>
<td>0.61</td>
</tr>
<tr>
<td>Apex Y/F Tuna Saku AAA</td>
<td>0.14</td>
<td>0.43</td>
</tr>
<tr>
<td>Blue Fin O-Toro</td>
<td>0.13</td>
<td>0.40</td>
</tr>
<tr>
<td>YF Tuna Loin</td>
<td>0.10</td>
<td>0.29</td>
</tr>
<tr>
<td>Bluefin Tuna Loin</td>
<td>0.08</td>
<td>0.25</td>
</tr>
<tr>
<td>BE Super Frozen Tuna Saku</td>
<td>0.03</td>
<td>0.10</td>
</tr>
<tr>
<td>Y/F. Tuna Loin</td>
<td>0.02</td>
<td>0.07</td>
</tr>
<tr>
<td>Chillian Sea Bass</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>BF Frozen O-Toro (Saku)</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>Awabi (Abalone)</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>Y/F Super Frozen Tuna Saku</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>Overall customized</td>
<td>32.87</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Note. Based on sales data and customization requests from a supplier of sushi restaurants in the southeastern United States, April 2010–February 2011.
example from a supplier of sushi restaurants. Online Appendix B contains formal proofs of all results.

2. Related Literature

There is a growing theoretical literature on relational contracts (Baker et al. 2002, Levin 2003, Gibbons 2005, Plambeck and Taylor 2006, Halac 2012, Li and Matouschek 2013, Halac 2015). Experimental and empirical research has also documented how relational contracts can ensure reliable product supply (Brown et al. 2004, Macchiavello and Morjaria 2015). A key finding of this literature is that discretionary payments can never become too large or the principal would violate the contract, and effort specified in the contract can never become too large or the agent would violate the contract. However, standard theoretical models of relational contracts assume that bonus payments are a general function of observed performance outcomes (e.g., Levin 2003). Such models do not allow meaningful analysis of cross-subsidized pricing. These models typically assume that there is only one task or product. Also, because they assume that payments can be a general function of outcomes for a task that requires effort, the principal does not pay a per-unit price premium for the task but instead offers a general bonus. By contrast, our model includes two products and a more realistic payment mechanism with constant per-unit prices that is commonly used in practice (Schmalensee 1989).

Consistent with our modeling assumptions, previous research has documented reasons why unit prices tend to remain constant even during demand fluctuations. Zbaracki et al. (2004) measure the costs of adjusting prices in a business-to-business (B2B) setting, finding large managerial costs of gathering information and communicating about price changes with the client. These costs of adjusting prices are a significant portion of the revenue and act as a deterrent toward changing prices too often or employing complicated pricing schemes. Moreover, Levy et al. (2010) show that prices remain rigid even during large demand spikes because of the opportunity cost of adjusting prices during these busy times. Laboratory experiments confirm that to avoid the implementation costs of dynamic pricing, subjects prefer to agree on a single price for multiple rounds of trading, even when doing so leads to some inefficient trades (Maciejovsky and Wernerfelt 2011). Previous research has also shown that fairness concerns can lead managers to set stable prices in order to avoid imposing disutility from unfairness on customers who might pay elevated prices during periods of high demand (Anderson and Simester 2010, Selove 2019).

Our paper uses an infinitely repeated game in the style of Rotemberg and Saloner (1986), Abreu (1988), Lal (1990), Dechenaux and Kovenock (2007), Thomadsen and Bhardwaj (2011), and Piccolo and Miklos-Thal (2012). Like these earlier papers, we derive conditions in which repeated interactions enable firms to cooperate in equilibrium. By applying this framework to the problem of pricing and quality provision in a channel with multiple products, we generate new insights about the optimal pricing tactics that make cooperation sustainable, including cross-subsidized pricing. Unlike papers on folk theorem results, we do not focus on cases in which the discount factors approach one, and therefore, it is not generally the case in our model that any outcome better than each firm’s worst-case payoff is sustainable. As is standard in the theory on relational contracts (e.g., Baker et al. 2002, Levin 2003) and most other theory papers on repeated interactions (e.g., Rotemberg and Saloner 1986), we derive the most efficient sustainable equilibrium for discount factors that do not approach one.

Previous theory on bundling and tying has developed models in which a monopolist requires customers who wish to purchase one product to also buy another product from the firm as an anticompetitive measure (Whinston 1990, Rey and Tirole 2007) or to extract more surplus from consumers (Bakos and Brynjolfsson 1999). Our model involves a different and, in some sense, opposite motivation for multiproduct relationships compared with this earlier literature. Rather than the supplier leveraging power over a price-taking customer to the detriment of the customer, firms in our model agree to prices that help to ensure that a long-standing and mutually valuable relationship persists while providing the strongest possible incentives for the supplier to exert effort that benefits the customer. Also, in our model, the supplier would not need to implement a formal legal requirement for customers to buy multiple products because the threat of reduced supplier effort provides an incentive for the downstream firm to buy multiple products from the supplier.

Literature on product line pricing has studied optimal prices when products are either complements or substitutes (Reibstein and Gatignon 1984, Dobson and Kalish 1988, Belloni et al. 2008). Our paper focuses on a different aspect of product line pricing. The goods in our model are neither substitutes nor complements on the demand side. Rather, firms agree on prices for multiple products that help sustain a reliable supply relationship with the greatest possible effort. Optimal prices in our model depend on the degree of volatility of demand for each good over time. If demand for the customizable good is sufficiently volatile, the downstream firm pays a premium both on that good and also on a basic good with more stable demand.

Bargaining as a mechanism of setting wholesale prices has been analyzed in Iyer and Villas-Boas (2003).
Our paper studies a different mechanism for sustaining effort, by allowing infinitely repeated interactions that make relational contracts possible. Another related stream of literature studies how information asymmetries affect channel partners (Jeuland and Shugan 1983, Desiraju and Moorthy 1997, Shaffer and Zettelmeyer 2002, Corbett et al. 2004, Busse et al. 2006, Guo and Iyer 2010). Our paper assumes that all parties in the channel have the same information about demand, costs and each other’s types.

The marketing literature also has a long tradition of studying relationships and their importance in B2B environments, studying characteristics of how such relationships influence transactions (Spekman et al. 1998, Jap 1999, Ghosh et al. 2006, Shervani et al. 2007, Tuli et al. 2007, 2010). Our paper contributes to this literature by formulating a theoretical model of repeated interactions that provides insights into the mechanism behind prices and quality in channel relationships.

3. Theoretical Model

The model has two risk-neutral players: a manufacturer and a retailer. The upstream firm can sell two goods to the downstream firm. Good 1 is a customizable good, the value of which increases if the seller invests (noncontractible) effort into customizing it for the retailer. Good 2 is a basic good that cannot be customized. The players trade repeatedly at dates $t = 0, 1, 2, \ldots$. Let good 1, which has an uncontractible quality dimension, be denoted by subscript $u$, whereas the basic good (good 2) is denoted by $b$.

On-time delivery, consistent good customer service, personal help, and even customized improvements to the physical good are possible examples of quality dimensions that may be uncontractable or unknown, as in Iyer and Villas-Boas (2003), at the time the bargaining process over prices takes place. The model allows for these uncontractable quality dimensions to be quite general; it suffices that a formal contract may be hard to enforce or hard to specify or costly to monitor for the results to be valid.

The firms find it impossible to write completely enforceable contracts because of the complexity of the transactions (for example), but their desire to continue to do business with each other in the future acts as an incentive to maintain a relational contract. Under certain conditions, this relationship is self-enforcing and leads to optimal quality for the buyer because both parties fear the loss of future benefits if they deviate from cooperating.

As is standard in the relational contracting literature (e.g., Baker et al. 2002, Levin 2003), including models of supply relationships (Macchiavello and Morjaria 2015, Andrews and Barron 2016), we abstract away from modeling the downstream firm’s customer demand. Rather, we assume that the retailer needs quantities $q_t,u$ and $q_t,b$ of goods 1 and 2, respectively, in period $t$. These amounts vary randomly from period to period, and we allow for $q_t,u$ and $q_t,b$ to be correlated in each period, with joint distribution $F(q_t,u, q_t,b)$ that is independent and identical across time periods. For now, the only other restriction we place on these distributions is that they are bounded, so there is a maximum possible quantity of each product. We later analyze the model for specific quantity distributions that allow us to explicitly derive optimal contracts that include a price premium on both products.

The downstream firm’s utility from buying from the seller at time $t$ is given by

$$U_{Rs} = q_t,u(\alpha_u + \beta_t u_t - P_t,u) + q_t,b(\alpha_b - P_t,b),$$

with $e_t$ the noncontractible effort per unit put in by the seller at time $t$, where $e_t \in [0, 1]$. Setting $e_t = 1$ implies that the supplier fully meets all the customization needs of the retailer for period $t$, whereas $e_t < 1$ implies that these needs are not fully met. Here $\beta$ is the per-unit discount of the customization effort to the retailer, $P_t,u$ and $P_t,b$ are the unit prices at time $t$, and $\alpha_u$ and $\alpha_b$ are constants.

The seller’s utility from selling to the retailer at time $t$ is given by

$$U_{Ss} = q_t,u(P_t,u - \gamma e_t) + q_t,b P_t,b,$$

with $\gamma$ being the per-unit cost of effort, where $\beta > \gamma > 0$, so the benefit to the retailer of customization effort exceeds the cost to the supplier. Note that optimal effort for maximizing total channel profits is given by $e_t = 1$ in each period. This is the optimal per-unit effort, which implies that optimal total effort costs for the seller are higher when quantity $q_t,u$ is high. We have assumed linear effort costs, similar to Bond and Gomes (2009) and Ludwig et al. (2011), so that we can explicitly derive the maximum sustainable effort and the efficient price levels for particular demand distributions.

Both firms have the option to buy or sell the goods on the outside market at unit prices $P_u$ and $P_b$, where $\alpha_u > P_u > 0$ and $\alpha_b > P_b > 0$. This outside option does not provide customization to the retailer and does not reward the supplier for the effort expended on customization. For example, the outside market may consist of low-end retailers who do not value customization and low-end suppliers who cannot serve customization requests in a cost-effective manner. Thus, for periods in which they trade on the outside market, the players utilities are given by

$$U_{Rs} = q_t,u(P_u - \bar{P}_u) + q_t,b(\alpha_b - \bar{P}_b),$$

$$U_{Ss} = q_t,u \bar{P}_u + q_t,b \bar{P}_b.$$
We later present a model extension that does allow the buyer to split its order for a period between different suppliers if it incurs an additional transaction cost.

The retailer wants to maximize \( \sum_{t=0}^{\infty} \delta^t U_{R_t} \), whereas the supplier wants to maximize \( \sum_{t=0}^{\infty} \delta^t U_{S_t} \), where each firm's discount factor is denoted by \( \delta \in (0, 1) \).

### 3.1. Timing

The parties initially agree on a pair of prices \( (P_u, P_b) \) for the goods. These prices can be greater than, equal to, or less than the outside market prices for the goods. The prices that the firms initially choose remain constant over the course of their relationship. The restriction to contracts with constant unit prices, which is commonly found in many industries, has been justified on the grounds of simplicity (Schmalensee 1989). One way to interpret the constant-unit-prices assumption is that there are often costs to adjusting unit prices (Levy et al. 1997, Zbaracki et al. 2004). The results in this paper show how firms can set prices of two different goods to achieve some flexibility in their pricing without the need to incur such price adjustment costs. We later present a model extension in which firms can agree on a different payment level for a given period than the one implied by their initial agreement if they incur an adjustment cost in that period.

After the players agree to prices for each good, they have the option to trade at times \( t = 0, 1, 2, \ldots \) based on this contract. At each time \( t \), the game has three stages:

1. Nature draws \( q_{t,u} \) and \( q_{t,b} \) and the buyer and the seller observe this demand for that particular period.
2. The buyer can either agree to pay the unit prices from the relational contract \( P_u \) and \( P_b \), or it can go to the outside market, which offers \( \bar{P}_u \) and \( \bar{P}_b \).
3. The seller decides how much effort to put into providing the service \( e_t \). The buyer observes this choice.

Note that the buyer commits to a payment before observing the seller’s effort level. Therefore, the seller has an opportunity to shirk on the contract by accepting payment for the goods while providing less effort than agreed by the relational contract. If effort levels were legally enforceable, then the parties could always achieve optimal effort with a formal contract, and a relational contract would not be necessary.

### 3.2. Results

A subgame perfect equilibrium, which we also refer to as a relationship or relational contract, consists of an initial pair of prices and strategies for each player after each possible history such that neither firm has a profitable deviation after any history. We define an equilibrium as efficient if it generates weakly greater total surplus than any other equilibrium, that is, if it maximizes the equilibrium value of the following:

\[
E \left( \sum_{t=0}^{\infty} \delta^t (U_{R_t} + U_{S_t}) \right).
\]

We will also say that one relational contract is more efficient than another if the former contract generates greater surplus than the latter. In principle, for any equilibrium that maximizes this surplus generated, firms could divide the surplus in any manner they choose with an initial lump-sum transfer between the players, as shown by Levin (2003).

In our model, all actions are observable by both players, so it is common knowledge whether either player has ever deviated from the equilibrium path. We will focus on equilibria in which, if either player deviates, firms never trade with each other again, which provides the strongest possible incentives for players to follow their equilibrium strategy (Abreu 1988, Levin 2003). In particular, because effort is observable in our model, temporary punishments are not necessary in equilibrium. In this respect, our model is similar to the repeated-game model with perfect monitoring by Rotemberg and Saloner (1986) and different from the models of imperfect monitoring by Green and Porter (1984) and Abreu et al. (1990).

Our formal propositions derive conditions in which it is efficient for firms to trade with each other during all periods. We first show that for such equilibria, it is efficient to set effort levels that depend only on the current quantity of the customizable good. Such relationships represent what Levin (2003) defines as a “stationary contract,” with effort and payment levels on the equilibrium path that depend only on the current period’s product quantities.

**Lemma 1.** If there exists an equilibrium in which firms always trade with each other, then the most efficient such equilibrium has effort levels \( e_t = e(q_{t,u}) \) that depend only on the current quantity of the customizable good.

We now derive conditions for both the supplier and the retailer to stay in a relationship in which they trade with each other during all periods.

At any time \( t \), the utility for the buyer if it stays in the relationship is given by its utility in the present period added to the discounted stream of utilities it gets in the future while it is in the relationship:

\[
q_{t,u}(a_u + \beta e_t - P_u) + q_{t,b} (a_b - P_b) + \sum_{j=t+1}^{\infty} \delta^{(T-j)} \times E[q_{T,u}(a_u + \beta e_T - P_u) + q_{T,b} (a_b - P_b)],
\]

where \( e_t \) and \( e_T \) are current and future supplier effort levels based on the relational contract.
By contrast, if the buyer deviates and goes to the outside market, the relationship will end. In this case, prices in this period and all future periods will be \( \tilde{P}_u \) and \( \tilde{P}_b \), and effort will be zero in the current period and all future periods. Thus, the buyer’s utility if it deviates at any time \( t \) is given by

\[
q_{t,u}(\alpha_u - \tilde{P}_u) + q_{t,b}(\alpha_b - \tilde{P}_b) + \sum_{t=1}^{\infty} \delta^{(t-t)} E[q_{T,u}(\alpha_u - \tilde{P}_u) + q_{T,b}(\alpha_b - \tilde{P}_b)]. \tag{7}
\]

By comparing expressions (6) and (7) and rearranging terms, we see that the buyer will always want to stay in the relationship if at each time \( t \):

\[
(P_u - \tilde{P}_u)\left(q_{t,u} + \frac{\delta}{1-\delta} E[q_{T,u}]\right) + (P_b - \tilde{P}_b)
\times \left(q_{t,b} + \frac{\delta}{1-\delta} E[q_{T,b}]\right) \leq \beta \left(q_{t,u} + \frac{\delta}{1-\delta} E[q_{T,u}]\right). \tag{8}
\]

Intuitively, the present value of all the premium payments the retailer makes must be less than the present value of the benefits it receives because of the supplier’s higher effort for the relationship to be sustainable.

The utility for the supplier if it stays in the relationship at any time \( t \) is

\[
q_{t,u}(P_u - \gamma e_t) + q_{t,b}P_b + \sum_{t=1}^{\infty} \delta^{(t-t)} E[q_{T,u}(P_u - \gamma e_t) + q_{T,b}P_b]. \tag{9}
\]

By contrast, if the supplier deviates and provides lower effort than specified in the relational contract, the relationship ends, and the supplier only receives prices \( \tilde{P}_u \) and \( \tilde{P}_b \) in all future periods. Therefore, at any time \( t \), the supplier’s utility, if it deviates and provides zero effort, is

\[
q_{t,u}P_u + q_{t,b}P_b + \sum_{t=1}^{\infty} \delta^{(t-t)} E[q_{T,u}\tilde{P}_u + q_{T,b}\tilde{P}_b]. \tag{10}
\]

By comparing expressions (9) and (10) and rearranging terms, we see that the supplier will always want to stay in the relationship if at each time \( t \)

\[
(P_u - \tilde{P}_u)\frac{\delta}{1-\delta} E[q_{T,u}] + (P_b - \tilde{P}_b)\frac{\delta}{1-\delta} E[q_{T,b}]
\geq \gamma \left(q_{t,u} + \frac{\delta}{1-\delta} E[q_{T,u}]\right). \tag{11}
\]

Intuitively, the present value of the premium payments the supplier receives in all future periods must be greater than the present value of its cost of effort in the current period and all future periods for the supplier to stay in the relationship.

We are interested in conditions in which firms can design a more efficient contract if they trade two goods with a price premium on both goods instead of placing a price premium only on the customizable good. In fact, we will now derive conditions in which there is no sustainable relationship in which firms always trade with each other with a price premium only on the customizable good. Formally, if we allow \( P_u - \tilde{P}_u > 0 \) but impose the constraint \( P_b - \tilde{P}_b = 0 \), then the buyer’s condition to stay in the relationship, given by (8), becomes

\[
(P_u - \tilde{P}_u)\left(q_{t,u} + \frac{\delta}{1-\delta} E[q_{T,u}]\right) \leq \beta \left(q_{t,u} + \frac{\delta}{1-\delta} E[q_{T,u}]\right). \tag{12}
\]

The seller’s incentive condition to stay in the relationship, given by (11), becomes

\[
\gamma \left(q_{t,u} + \frac{\delta}{1-\delta} E[q_{T,u}]\right) \leq (P_u - \tilde{P}_u)\frac{\delta}{1-\delta} E[q_{T,u}]. \tag{13}
\]

Multiplying both sides of (12) by \( \frac{\gamma}{\beta} \) and both sides of (13) by \( \frac{\beta}{\gamma^2} \) and combining these two constraints imply that the following must hold:

\[
\gamma \left(q_{t,u} + \frac{\delta}{1-\delta} E[q_{T,u}]\right) \leq \beta \frac{\delta}{1-\delta} E[q_{T,u}]. \tag{14}
\]

Rearranging terms shows that this condition cannot hold during periods of peak demand of the customizable good, that is, when \( q_{t,u} = \max(q_{t,u}) \), if the following is true:

\[
\gamma \max(q_{t,u}) > (\beta - \gamma) \frac{\delta}{1-\delta} E[q_{T,u}]. \tag{15}
\]

Therefore, if (15) holds, then it is not possible for both (8) and (11) to hold during all periods for any relational contract that has a price premium on the customizable good but not on the basic good.

Intuitively, if (15) holds, then demand spikes for the customizable good are so large that it is not possible to sustain optimal effort (\( e_t = 1 \)) during periods with the highest demand for this good. Therefore, the premium payment during such periods must be lowered to reflect the seller’s actual effort level during peak demand to induce the buyer to make this payment. However, if we constrain the contract to include a price premium only on the customizable good (and not on the basic good), then reducing the payment during peak demand proportionally reduces the payment level during other periods, resulting in even weaker effort incentives, which require an additional payment reduction during peak demand to reflect the lowered effort levels, and so on, until any potential contract unravels.
Proposition 1. If inequality (15) holds, then it is not possible to sustain any relationship in which firms always trade with each other with a price premium only on the customizable good. By contrast, if inequality (15) is reversed, firms can sustain optimal effort \( (e_i = 1) \) during all periods with a price premium only on the customizable good.

The next step in the analysis is to identify an arrangement that gives both the seller and the buyer the appropriate incentives to sustain a relational contract during all periods when condition (15) holds. We will now derive an efficient contract that includes premium payments on both goods instead of only on the customizable good. To illustrate how this can occur, and to allow us to explicitly derive the most efficient contract, we assume that the quantities traded for each product have the following distribution:

\[
\begin{align*}
\text{Prob}(q_{i,a} = L, q_{i,b} = L) &= 1 - \omega, \\
\text{Prob}(q_{i,a} = H_a, q_{i,b} = H_b) &= \omega, 
\end{align*}
\]

where \( 0 < L < H_b < H_a \) and \( \omega \in (0, 1) \). Note that we are assuming that demand for the goods is perfectly correlated and that there are only two possible states of demand in each period: one state in which both goods have low demand and another state in which both goods have relatively high demand. In the next two sections of this paper, we present extensions that allow demand for the two goods to be uncorrelated or imperfectly correlated. For notational simplicity, we have also normalized demand for the two goods to be equal \( (q_{i,a} = q_{i,b} = L) \) in the low condition. However, this assumption could be relaxed, and the prices for the basic good could be scaled accordingly to produce equivalent results.

We assume that the high-demand quantity for the customizable good is large enough to satisfy the following:

\[
\gamma H_a > (\beta - \gamma) \left( \frac{\delta}{1 - \delta} \right) [(1 - \omega)L + \omega H_a].
\]  

Note that (18) is equivalent to (15) for this demand distribution. If inequality (18) were reversed, then the second part of Proposition 1 would imply that firms can sustain an efficient relationship with a price premium only on the customizable good.

We also assume that firms are patient enough that the following holds:

\[
\gamma < (\beta - \gamma) \left( \frac{\delta}{1 - \delta} \right).
\]  

If inequality (19) were reversed, then firms could not sustain any relationship, even if the quantity for the customizable good were constant.

When inequalities (18) and (19) both hold, we show that the most efficient relationship includes a price premium on both goods, and this relationship can be sustained during all periods. We will now derive the most efficient contract. This contract involves setting \( e_i = 1 \) during low demand and setting the maximum sustainable effort, denoted \( e^*_i \), during high demand. We will show that the optimal high-demand effort level can be found by solving the following equation:

\[
\gamma H_a e^*_i = (\beta - \gamma) \left( \frac{\delta}{1 - \delta} \right) [(1 - \omega)L + \omega H_a e^*_i].
\]  

When this equation holds, effort costs during high demand equal the expected discounted value of the relationship, so this effort level is sustainable. Note that each side of (20) is a linear function of \( e^*_i \). Furthermore, (19) implies that the left side of (20) is less than the right side if we set \( e^*_i = \frac{L}{H_a} \), whereas (18) implies that the left side of (20) is greater than the right side if we set \( e^*_i = 1 \). Therefore, there is a unique value of \( e^*_i \in (\frac{L}{H_a}, 1) \) that satisfies (20). By solving (20), we find that

\[
e^*_i = \frac{L}{H_a} \left[ \frac{(\beta - \gamma) \left( \frac{\delta}{1 - \delta} \right) [(1 - \omega)L + \omega H_a]}{\gamma - (\beta - \gamma) \left( \frac{\delta}{1 - \delta} \right) \omega} \right].
\]

Equation (21) gives the maximum sustainable per-unit effort level during high demand. Note that \( e^*_i \) is increasing in the marginal benefit of effort \( (\beta) \) and the weight that firms place on the future (represented by \( \delta) \), but decreasing in the marginal cost of effort \( (\gamma) \) and the relative magnitude of high-demand quantity of the customizable good \( (H_a) \).

Given two prices and two possible demand states, the firms can design a contract with the precise total payment it would like to set in each demand state. We show that the relational contract can sustain effort \( e_i = 1 \) during low demand and \( e_i = e^*_i \) during high demand by setting the following prices:

\[
P_{u} - \bar{P}_{u} = \frac{(H_a e^*_i - H_b)}{(H_a - H_b) \beta},
\]

\[
P_{b} - \bar{P}_{b} = \beta - (P_{u} - \bar{P}_{u}).
\]

As the maximum sustainable effort during high demand \( (e^*_i) \) falls, the contract shifts the price premium toward the basic good; that is, \( P_{u} - \bar{P}_{u} \) decreases, and \( P_{b} - \bar{P}_{b} \) increases by the same amount. As \( e^*_i \to \frac{L}{H_a} \), the contract converges to placing the entire premium payment on the basic good. By contrast, as \( e^*_i \to 1 \), the contract converges to placing the entire premium payment on the good with uncontractible quality.

Proposition 2. If the quantity distributions satisfy conditions (18) and (19), then the most efficient relational contract includes a price premium on the basic good, with low-demand effort level \( e_i = 1 \), high-demand effort level given by (21), and prices given by (22) and (23).
Intuitively, including a price premium on the more stable, noncustomized good provides larger total premium payments to the supplier in times of high demand without letting the total premium payment grow so large that the retailer would want to exit the relationship. This same contract is not sustainable if the same premium payment is placed only on the good with uncontractible quality because the premiums become more than the retailer is willing to pay during the large spikes in quantity traded of this good.

Proposition 2 implies that if demand for the basic good is volatile enough that the following condition holds, then the efficient contract sets a negative price premium on the customizable good:

$$\gamma H_b > \left(\beta - \gamma \right) \left(\frac{\delta}{1 - \delta}\right) [(1 - \omega)L + \omega H_b].$$

(24)

Recall that $H_u > H_b$, which implies that demand for the customizable goods is more volatile than for the basic good even when (24) holds. In this case, both goods have such volatile demand that a price premium on just one good could not be sustained, and the efficient contract involves a large positive premium for the relatively less volatile (basic) good and a small negative premium for the more volatile (customizable) good. This contract matches the optimal payment level during both low- and high-demand states. During high demand, the small negative price premium on the customizable good helps offset the increased premium payment that results from greater demand for the basic good.

**Corollary 1.** If the conditions of Proposition 2 hold and demand for the basic good is volatile enough that (24) holds, then the efficient contract includes a positive price premium for the basic good ($P_b > \bar{P}_b$) and a negative price premium for the customizable good ($P_u < \bar{P}_u$).

Finally, we show that Proposition 2 implies that the optimal contract converges to a contract with a price premium only on the basic good (with no price premium on the customizable good) as the magnitude of demand spikes for the customizable good grows.

**Corollary 2.** Starting with the conditions of Proposition 2 and letting the magnitude of demand spikes for the customizable good grow ($H_u \to \infty$), the optimal price premium on the customizable good approaches zero (($P_u - \bar{P}_u) \to 0$), and the optimal price premium on the basic good approaches the marginal benefit of effort (($P_b - \bar{P}_b) \to \beta$).

As is standard in the theoretical contracting literature, we have focused on the most efficient contract in terms of total surplus generated, without analyzing how that surplus is shared between the players (see Weitzman 1980, Bolton and Dewatripont 2005). In fact, in our model, all the surplus goes to the supplier in order to provide the strongest possible incentives for supplier effort during high demand. In principle, firms could split the surplus differently by using an initial lump-sum transfer. For example, the supplier could offer the retailer a one-time new customer discount on the first payment as they begin their relationship. Because this initial payment occurs before any effort by the supplier, such a discount would not affect the supplier’s effort incentives. Therefore, this slightly modified version of the relationship we have derived, with an initial lump-sum transfer to the buyer, would also be sustainable.

### 3.3. Numerical Example

We now present a numerical example to help provide more intuition for the results from the preceding section. Table 2 presents the parameter values used for the numerical example.

For this example, firms experience occasional demand shocks: The parameter $\omega = 0.1$ indicates that 90% of time periods have low demand, whereas 10% have high demand. We normalize low-demand quantity of each good to five units ($L = 5$), and we allow demand for the basic good to double during high demand ($H_b = 10$), while demand for the customizable good increases by an even greater amount ($H_u \in [15, 40]$). The per-unit cost of effort is normalized to $\gamma = 1$, and the per-unit benefit of effort is greater than the cost ($\beta \in [1.1, 2.1]$). Firms are patient enough to care about future periods but not infinitely patient, so we set $\delta = 0.8$.

Figure 1 presents the optimal price policy as a function of the per-unit benefit of effort and the high-demand quantity of the customizable good. Holding other parameters constant, an increase in the per-unit benefit of effort ($\beta$) implies that the relationship

<table>
<thead>
<tr>
<th>Table 2. Parameter Values Used in the Numerical Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Value</strong></td>
</tr>
<tr>
<td>$\omega = 0.1$</td>
</tr>
<tr>
<td>$L = 5$</td>
</tr>
<tr>
<td>$H_b = 10$</td>
</tr>
<tr>
<td>$H_u \in [15, 40]$</td>
</tr>
<tr>
<td>$\gamma = 1$</td>
</tr>
<tr>
<td>$\beta \in [1.1, 2.1]$</td>
</tr>
<tr>
<td>$\delta = 0.8$</td>
</tr>
</tbody>
</table>
becomes more valuable, and therefore, high payment and high effort levels are easier to sustain. By contrast, an increase in the high-demand quantity of the customizable good \((H_u)\) implies that premium payments on the customizable good have a larger increase during high demand, and therefore, the firms may need to shift part of the premium payment to the basic good so that the total premium payment is always sustainable.

If the per-unit benefit of effort is too low \((\beta < 1.25)\), it is not possible to sustain any relationship. Formally, if \(\beta\) is small enough that \(\gamma > (\beta - \gamma)\left(\frac{1}{1 - \beta}\right)\), then condition (19) does not hold, and the average effort cost in any possible relationship would exceed the value of maintaining the relationship, so there is no payment structure that can sustain any relationship.

For slightly higher values of the benefit of effort \((\beta \in [1.25, 1.45])\), firms can sustain a relationship, but this relationship is not valuable enough to sustain effort cost \(\gamma H_b\) during high demand. In this case, a per-unit premium of magnitude \(\beta\) on the basic good, with no premium on the customizable good, would lead to a payment spike during high demand that is too large to sustain. Therefore, firms place a small negative price premium on the customizable good to offset the payment spike on the basic good. Formally, under the conditions of Proposition 2, and given \(\beta\) small enough that (24) holds, Corollary 1 implies that firms place a positive price premium on the basic good but a negative premium on the customizable good.

For higher values of the benefit of effort \((\beta > 1.45)\), the optimal price strategy depends on the magnitude of the high-demand quantity of the customizable good \((H_u)\). If these demand spikes are large enough that (15) holds, Proposition 2 implies that firms place a positive price premium on both goods. By contrast, if \(H_u\) is small enough that (15) does not hold, Proposition 1 implies that firms place a price premium only on the customizable good.\(^{12}\)

### 3.4. Model Extension: Payment Adjustment Costs

In previous sections, firms agreed to a relational contract with constant unit prices. Ideally, firms would like to use such a contract, or other simple contracts, to generate the optimal premium payment for every possible state of demand. However, if the number of possible states of demand is too large, it may be too complicated to construct a contract that specifies the optimal payment for every possible state of demand. In such cases, firms could agree to a contract that generates optimal payments for the most likely states of demand while also agreeing to negotiate a different premium payment than the one specified by the contract during periods when relatively unlikely states occur. In this section, we extend our model to allow firms to negotiate a different premium payment than the one implied by the unit prices in their contract, if they incur an adjustment cost for the period.

For example, suppose that firms agree to a contract that includes a premium payment on both a basic good and a customizable good. This contract may yield optimal payments when the quantity of both goods is low and when the quantity of both goods is relatively high. However, the firms may want to negotiate a lower total payment than the one implied by these prices during periods when demand for the basic good is high but demand for the customizable good is low in order to reflect the relatively low value of effort during such periods. Similarly, they may want to negotiate a higher total payment than the one implied by the contract’s prices during periods when demand for the basic good is low but demand for the customizable good is high.

More generally, if there are many possible states of demand, and it is too complicated to construct a contract that specifies the optimal payment for every possible demand state, firms could set prices that generate the optimal payment for the most likely demand states while also reaching an understanding of the types of situations in which they will deviate from these prices. Formally, the relationship derived in this model extension represents an incomplete contract (Hart and Moore 1999).

We assume that firms initially agree to a relational contract that specifies values \((P_{u1}, P_b)\) representing unit prices for each good. However, in each period, we allow firms to agree to a different payment than the one implied by these prices, similar to the model of procurement contract renegotiation by Bajari and Tadelis (2001). In principle, we could allow firms to negotiate new prices that continue to hold after the adjustment. However, it is more efficient to negotiate a one-time payment adjustment and then revert to the
original contract prices if the original prices were chosen optimally for the given demand distributions.

The timing in each period is the following:
1. Nature draws \( q_{t,u} \) and \( q_{t,b} \) and the buyer and the seller observe this demand for that particular period.
2. The seller has the option, if it incurs an adjustment cost \( d \), of proposing an alternative bonus payment \( B_t \) (in excess of what the buyer would pay on the outside market) for the current period, which can be different from the bonus implied by the prices in the relational contract.
3. If the seller has proposed such an alternative payment, the buyer can either agree to make this proposed payment or can go to the outside market. If the seller has not proposed an alternative payment in the current period, the buyer can either agree to pay the unit prices specified in the relational contract or can go to the outside market.
4. The seller decides how much effort \( e_t \) to put into providing the service. The buyer observes this choice.

When the seller proposes and the buyer agrees to an alternative bonus \( B_t \), the buyer’s utility for the period is
\[
U_{B_t} = q_{t,u}(\alpha_u + \beta e_t - \bar{p}_u) + q_{t,b}(\alpha_b - \bar{p}_b) - B_t,
\]
and the seller’s utility for the period is
\[
U_{S_t} = q_{t,u}(\bar{p}_u - \gamma e_t) + q_{t,b} \bar{p}_b + B_t - d,
\]
where the adjustment cost for the period is represented by \( d > 0 \).

We characterize subgame perfect equilibria based on contract prices \( (P_u, P_b) \), an effort level for each demand state, a set of demand states \( D \) for which the seller will propose an alternative bonus, and the alternative bonus payment \( B(q_{t,u}, q_{t,b}) \) for each demand state \( (q_{t,u}, q_{t,b}) \in D \). As before, there is a bad equilibrium in which the buyer always goes to the outside market and the seller always invests zero effort, and firms revert to this bad equilibrium if either party ever deviates from the proposed equilibrium.

We now generalize the demand distribution used in the final proposition of the preceding section, allowing for imperfect correlation in product quantities. In particular, we assume that the quantities traded for each product have the following distribution:
\[
\begin{align*}
\text{Prob}(q_{t,u} = L, q_{t,b} = L) &= (1 - \rho)(1 - \omega)^2 + \rho(1 - \omega), \\
\text{Prob}(q_{t,u} = L, q_{t,b} = H_b) &= (1 - \rho)\omega(1 - \omega), \\
\text{Prob}(q_{t,u} = H_u, q_{t,b} = L) &= (1 - \rho)\omega(1 - \omega), \\
\text{Prob}(q_{t,u} = H_u, q_{t,b} = H_b) &= (1 - \rho)\omega^2 + \rho \omega,
\end{align*}
\]
where \( 0 < L < H_b < H_u, \omega \in (0,0.5) \), and \( \rho \in [0,1] \). For these distributions, each product has probability \((1 - \omega)\) of having a low quantity and probability \( \omega \) of having a relatively high quantity. As in the preceding section, we assume that the high-demand quantity for the customizable good is large enough that (18) holds and that firms are patient enough that (19) holds.

The parameter \( \rho \) reflects their strength of correlation, with \( \rho = 0 \) implying no correlation and \( \rho = 1 \) implying perfect correlation. If the correlation in product demands is sufficiently strong, then the two most likely states are that both products have low demand and that both products have relatively high demand. In particular, states \((L,L)\) and \((H_u,H_b)\) are each more likely than state \((H_u,L)\) and more likely than state \((L,H_b)\) if the following holds:
\[
\rho > \max\left(-\frac{1 + 2\omega}{2\omega^2}, 1 - \frac{2\omega}{2 - 2\omega}\right). \tag{31}
\]

If the correlation in product quantities is strong enough that this condition holds, an efficient contract uses unit prices to generate optimal payments at states \((L,L)\) and \((H_u,H_b)\), and firms make price adjustments at states \((H_u,L)\) and \((L,H_b)\). Furthermore, we show that such a contract involves premium payments on both the customizable good and the basic good.

The derivations of this contract are similar to those for Proposition 2. Condition (18) implies that effort \( e_t = 1 \) cannot be sustained during high demand. If the adjustment cost \( d \) is sufficiently small, then the optimal contract involves setting \( e_t = 1 \) during low demand and \( e_t = e^*_H \) during high demand, where the maximum sustainable effort during high demand is found by solving the following:
\[
\gamma H_e^* = \left(\frac{\delta}{1 - \delta}\right)[(\beta - \gamma)[(1 - \omega)L + \omega H_u e^*_H] - 2d(1 - \rho)x(1 - \omega)], \tag{32}
\]
where the term \( 2d(1 - \rho)x(1 - \omega) \) reflects expected adjustment costs in each period, that is, the adjustment cost \( d \) times the probability that one product has low demand while the other has relatively high demand. For sufficiently small \( d \), conditions (18) and (19) ensure that there is a value \( e^*_H \in \left(\frac{\delta}{1 - \delta}, 1\right) \) that solves this equation. Solving for \( e^*_H \), we find the following:
\[
e^*_H = \frac{1}{H_u} \left[ \frac{\beta - \gamma (1 - \omega)L - 2d(1 - \rho)x(1 - \omega)}{\gamma - \frac{\delta}{1 - \delta}(\beta - \gamma)x} \right]. \tag{33}
\]

Note that the maximum sustainable effort during high demand is decreasing in the adjustment cost \( d \) and increasing the strength of product quantity correlation \( \rho \).
As in Proposition 2, optimal prices as a function of \( c_{it}^* \) are given by the following:

\[ P_u - \bar{P}_u = \left( \frac{H_u c_{it}^* - H_b}{H_u - H_b} \right) \beta, \]  
\[ P_b - \bar{P}_b = \beta - (P_u - \bar{P}_u). \]  
\[ (34) \quad (35) \]

Firms make payment adjustments at states \((H_u, L)\) and \((L, H_b)\), and the associated bonus payments are

\[ B(H_u, L) = \beta H_u c_{it}^*, \]  
\[ B(L, H_b) = \beta L. \]  
\[ (36) \quad (37) \]

**Proposition 3.** If \( d \) is sufficiently small and the quantity distributions satisfy conditions (18), (19), and (31), then the most efficient relational contract includes a premium payment on the basic good, with low-demand effort level \( e_t = 1 \), high-demand effort level given by (33), unit prices given by (34) and (35), and payment adjustments given by (36) and (37).

Thus, if there is a sufficiently strong correlation in demand for a customizable good with very volatile demand and a basic good with less volatile demand, including a premium payment on both goods can generate an optimal payment when demand for both goods is low and when demand for both goods is relatively high. The firms can then make costly payment adjustments when other less likely states occur.

The stylized example in this section illustrates the more general point that when the number of possible demand states is too large for the contract to specify a payment in every possible state, firms can use unit prices on multiple goods to achieve some pricing flexibility with a relatively simply contract.

### 3.5. Model Extension: Order Splitting

In this section, we allow the buyer to split its order in a given period between the supplier and the outside market if it incurs a transaction cost for the period. For example, suppose that the buyer and supplier agree to a contract that includes a premium payment on the customizable good. During periods in which demand for this good is particularly high, the buyer can buy a fraction of its demand for the good from the outside market to limit its total premium payment for the period. Previous theoretical research on relational contracts by Board (2011) and Andrews and Barron (2016) shows that using multiple suppliers can reduce their effort incentives because each supplier receives a smaller share of future premium payments. By contrast, in our model with fixed linear prices, we show that order splitting during high demand can actually help firms sustain a relationship with high effort by limiting premium payments to a level that is acceptable to the buyer.

We assume that firms initially agree to a relational contract that specifies values \((P_u, P_b)\), representing unit prices for each good. The timing in each period is the following:

1. Nature draws \( q_{t,u} \) and \( q_{t,b} \), and the buyer and the seller observe this demand for that particular period.
2. As before, the buyer can buy all its goods from the seller at the prices specified in the contract or can buy all its goods from the outside market. Alternatively, if it incurs transaction cost \( f \), the buyer can buy a fraction \( F_{t,u} \) of the customizable good and a fraction \( F_{t,b} \) of the basic good from the seller at the prices specified in the contract and buy the remaining fractions \( 1 - F_{t,u} \) and \( 1 - F_{t,b} \), respectively, from the outside market, where the buyer can choose any \( F_{t,u} \in [0,1] \) and \( F_{t,b} \in [0,1] \).
3. The seller decides how much effort \( e_t \) to put into providing the service. This effort applies only to the quantity of the customizable good that the buyer purchases from the seller, not from the outside market. The buyer observes this choice.

When the buyer splits its order, the buyer’s utility for the period is

\[ U_b = q_{t,u} \left[ \alpha_u + F_{t,u} \beta e_t - F_{t,u} \bar{P}_u - (1 - F_{t,u}) \bar{P}_b \right] 
+ q_{t,b} \left[ \alpha_b - F_{t,b} \bar{P}_b -(1 - F_{t,b}) \bar{P}_b \right] - f. \]  
\[ (38) \]

And the seller’s utility for the period is

\[ U_s = q_{t,u} \left[ F_{t,u} (P_u - \gamma e_t) + (1 - F_{t,u}) \bar{P}_u \right] 
+ q_{t,b} \left[ F_{t,b} \bar{P}_b + (1 - F_{t,b}) \bar{P}_b \right]. \]  
\[ (39) \]

We characterize subgame perfect equilibria based on contract prices \((P_u, P_b)\), an effort level for each demand state, and the buyer’s purchase decision (including possible order splitting) for each demand state. Firms revert to the bad equilibrium, in which the two parties never trade with each other again, if either party ever deviates from the proposed equilibrium.

As in the preceding section, the quantities traded for the two goods have joint distribution given by (27)–(30), and we assume that conditions (18) and (19) hold. We also assume that demand for the basic product is not too volatile, so condition (24) does not hold, which implies that the optimal contract derived in previous sections involves a positive price premium on both goods. In this case, order splitting does not allow the same flexibility as the payment adjustments modeled in the preceding section. Payment adjustments allow the firm to increase or decrease the premium payment in any given period. By contrast, order splitting allows the firm to decrease its quantity of each good purchased from the supplier in a period but does not allow for an increase in quantity demanded. Therefore, a relationship such as the one derived in the preceding section, with payment adjustments only when one product has high demand but the other has low demand, cannot be sustained using order splitting.
because such a relationship requires the flexibility for both upward and downward adjustments.\textsuperscript{15}

We show that the optimal contract with order splitting does not involve a price premium on both goods. Rather, the optimal contract involves a price premium only on the customizable good, with order splitting during high demand for this good and effort \( e_1 = 1 \) during all periods. The following equation states the buyer’s constraint for such a contract:

\[
(\beta - P_u + \bar{P}_u) \left[ H_u F_{H,u} + \left( \frac{\delta}{1 - \delta} \right) (\alpha H_u F_{H,u}) + (1 - \omega)L \right] - f \left[ 1 + \left( \frac{\delta}{1 - \delta} \right) \omega \right] \geq 0,
\]

where \( F_{H,u} \) denotes the share of the customizable good purchased from the supplier during high demand. The first term on the left-hand side of (40) represents the value of effort minus the price premium multiplied by the high-demand quantity plus the expected value of all future quantities of the customizable good. The second term reflects the order-splitting cost during high demand plus the expected value of all future order-splitting costs. The maximum price for which this constraint holds is given by

\[
P_u - \bar{P}_u = \beta - \phi f, \tag{41}
\]

where \( \phi \) is defined as:

\[
\phi = \frac{1 + \left( \frac{\delta}{1 - \delta} \right) \omega}{H_u F_{H,u} + \left( \frac{\delta}{1 - \delta} \right) (\alpha H_u F_{H,u}) + (1 - \omega)L}. \tag{42}
\]

If order-splitting costs \( f \) are sufficiently small, we show that the optimal contract involves a price premium only on the customizable good, with order splitting during high demand for this good and effort \( e_1 = 1 \) during all periods. The fraction of the customizable good purchased from the supplier during high demand is found by solving the following:

\[
y H_u F_{H,u} = \left( \frac{\delta}{1 - \delta} \right) \left[ (\beta - \gamma - \phi f) \times [ (1 - \omega)L + \alpha H_u F_{H,u}] \right]. \tag{43}
\]

Inequality (18) ensures that the left-hand side of this equation is greater than the right-hand side for \( F_{H,u} = 1 \), whereas (19) ensures that for sufficiently small \( f \), the left-hand side is less than the right-hand side for \( F_{H,u} = \frac{1}{\gamma} \). Because both sides are continuous functions of \( F_{H,u} \), there must be a value \( F_{H,u} \in (\frac{1}{\gamma}, 1) \) that solves this equation.\textsuperscript{16}

**Proposition 4.** If order-splitting cost \( f \) is sufficiently small and the quantity distributions satisfy (18) and (19) but not (24), then the efficient relationship sets effort \( e_1 = 1 \) during all periods, with no order splitting during low demand, order splitting during high demand given by solving (43) for \( F_{H,u} \), and a price premium only on the customizable good \( P_u - \bar{P}_u = \beta - \phi f \).

The following corollary compares the relative attractiveness of order splitting and payment adjustments.

**Corollary 3.** If the conditions of Propositions 3 and 4 hold, then payment adjustments are more efficient than order splitting if \( 2\delta(1 - \rho)(1 - \omega) < f \).

Firms prefer a relationship with payment adjustments if the correlation in product demands \( (\rho) \) is high, in which case payment adjustments are rarely needed because most periods involve either high demand for both goods or low demand for both goods. Firms also prefer payment adjustments if the probability of high demand \( (\omega) \) is high, in which case they would like to avoid the need for order splitting during all the periods with high demand for the customizable good. By contrast, firms tend to prefer a relationship with order splitting if the correlation between product demands is low, the probability of high demand is low, and the cost of order splitting \( (f) \) is less than the cost of payment adjustment \( (d) \).

### 3.6. Model Extension: Two-Part Tariffs

We now extend the model to allow contracts with a fixed-payment component and per-unit prices, similar to two-part tariffs. Although including a fixed payment improves contract efficiency, we show that fixed payments do not generally eliminate the benefit of setting a per-unit price premium on both the basic good and the customizable good. Firms can still use per-unit price premiums on multiple goods to provide additional degrees of freedom in determining the premium payment in each period. In the example in this section, setting a price premium on both goods allows for payments that increase up to a certain quantity of the customizable good and then become less sensitive to quantity above that level.

Formally, we allow firms to agree to a relational contract that specifies values \( (y, P_u, P_b) \), where \( y \) is the fixed-payment component, which can be either positive or negative (or zero). When the buyer agrees to the payment specified in the relational contract, the buyer’s utility for the period is

\[
U_r = q_{i,u}(\alpha + \beta e_1 - P_u) + q_{i,b}(\alpha_b - P_b) - y. \tag{44}
\]

And the seller’s utility for the period is

\[
U_s = q_{i,u}(P_u - \gamma e_1) + q_{i,b}P_b + y. \tag{45}
\]

Using a fixed payment and per-unit prices on two products allows three degrees of freedom, which generally allows firms to generate the exact premium payment they would like in three different demand states. To derive an efficient contract for this case, we extend the demand distribution used in Proposition 2 to include three possible demand states. In particular,
we assume that the quantities traded for each product have the following distribution:

\[
\begin{align*}
\text{Prob}(q_{i,M} = L_u, q_{i,b} = L_b) &= 1 - \omega_M - \omega_H, \\
\text{Prob}(q_{i,M} = M_u, q_{i,b} = L_b + M_u - L_u) &= \omega_M, \\
\text{Prob}(q_{i,M} = H_u, q_{i,b} = L_b + M_u - L_u) &= \omega_H,
\end{align*}
\] (46)

(47)

(48)

where \(0 < L_u < M_u < H_u, L_b > 0, \omega_M > 0, \omega_H > 0\), and \(\omega_M + \omega_H < 1\). This distribution implies that both products exhibit an increase in quantity of \(M_u - L_u\) as they move from the low-demand state to the medium-demand state; however, the customizable good has a further increase in quantity of \(H_u - M_u\), whereas the basic good has no further increase in quantity as they move from the medium-demand state to the high-demand state. Other distributions with three states could be used to produce similar results, with appropriate adjustments to the fixed payment and per-unit prices, as long as each product’s quantity is not a perfectly linear function of the other product’s quantity.

For simplicity of exposition, we have returned to the assumption of the original model with perfect correlation in demand between the products. In principle, one could allow imperfect correlation between the products, with payment adjustments or order splitting to match the desired payment during some demand states, as in the preceding model extensions.\(^{17}\)

We focus on cases in which these distributions satisfy the following conditions:

\[
\begin{align*}
\gamma H_u > (\beta - \gamma) \left( \frac{\delta}{1 - \delta} \right) [(1 - \omega_M - \omega_H) L_u \\
+ \omega_M M_u + \omega_H H_u], \quad \text{(49)}
\end{align*}
\]

\[
\begin{align*}
\gamma M_u < (\beta - \gamma) \left( \frac{\delta}{1 - \delta} \right) [(1 - \omega_M - \omega_H) L_u \\
+ (\omega_M + \omega_H) M_u]. \quad \text{(50)}
\end{align*}
\]

Condition (49) implies that effort \(e_i = 1\) cannot be sustained during high demand, but condition (50) implies that effort \(e_i = 1\) can be sustained during both low and medium demand. The optimal contract involves setting \(e_i = 1\) during low and medium demand and setting \(e_i = e^*_H\) during high demand, where the maximum sustainable effort during high demand is found by solving the following:

\[
\begin{align*}
\gamma H_u e^*_H = (\beta - \gamma) \left( \frac{\delta}{1 - \delta} \right) [(1 - \omega_M - \omega_H) L_u \\
+ \omega_M M_u + \omega_H H_u e^*_H]. \quad \text{(51)}
\end{align*}
\]

Conditions (49) and (50) ensure that there is a value \(e^*_H \in \left( \frac{\omega_H}{H_u}, 1 \right) \) that solves this equation. Solving for \(e^*_H\), we find

\[
e^*_H = \frac{1}{H_u} \left( \frac{\beta - \gamma}{\gamma - (\beta - \gamma) \frac{\omega_H}{H_u}} \right) \left[ (1 - \omega_M - \omega_H) L_u + \omega_M M_u \right]. \quad \text{(52)}
\]

The optimal per-unit prices and fixed payment that sustain these effort levels are as follows:

\[
\begin{align*}
P_u - \bar{P}_u &= \left( \frac{H_u e^*_H - M_u}{H_u - M_u} \right) \beta, \quad \text{(53)}
\end{align*}
\]

\[
\begin{align*}
P_b - \bar{P}_b &= \beta - (P_u - \bar{P}_u), \quad \text{(54)}
\end{align*}
\]

\[
y = \beta L_u - (P_u - \bar{P}_u) L_u - (P_b - \bar{P}_b) L_b. \quad \text{(55)}
\]

**Proposition 5.** If the quantity distributions satisfy conditions (49) and (50), then the most efficient relational contract includes a fixed payment and per-unit premium payments on both goods, with low- and medium-demand effort levels \(e_i = 1\), high-demand effort level given by (52), unit prices given by (53) and (54), and the fixed payment given by (55).

Thus, even when the relational contract includes a fixed-price component, it is generally optimal to include a per-unit premium payment on both goods to achieve additional pricing flexibility.

4. Conclusion

We develop a model in which a downstream firm pays premium prices on the goods it purchases in order to compensate its supplier for customization services that are not enforceable through a formal contract. We show that the payments that can sustain such a relationship are partially linked to goods with stable demand, even if these goods are not customizable. Moreover, this relational contract ensures efficient provision of uncontractible services for the customizable goods.

Our results provide one explanation for why firms often use cross-subsidized pricing. In some cases, it is efficient for a firm to pay a price premium on one good to compensate a supplier for effort expended on another good. We show that this pricing strategy can help channel members form a simple relational contract that provides price flexibility to compensate suppliers for greater effort during high demand without making the total premium payment too large to sustain.

In particular, if marginal effort by the supplier generates a moderate amount of surplus (surplus generated is not too low or too high), then firms can sustain a relationship, but they cannot sustain maximum effort during large demand spikes. In such cases, a price premium only on the customizable good would not be sustainable because it would lead to payment spikes that exceed the value of effort during high demand. We derive conditions in which the most efficient relationship sets a price premium on multiple goods so that the total premium payment during each possible demand state equals the value of sustainable effort in that demand state. As the volatility of the customizable good increases, firms need to shift the price premium away from this volatile good and toward a more stable good to maintain a given
difference in the total premium payment between high- and low-demand states.

We do not explicitly model customer demand and retail prices, which could lead to double marginalization. Further research could explore relational contracts as a tool to solve the problem of double marginalization. In addition, it would be interesting to explore what happens if the retailer also can incur uncontractible effort that increases sales and benefits the channel, so the contract needs to provide effort incentives for both firms.

Future research could also model how firms initially bargain over the terms of a relational contract. In our model, formulating an efficient contract requires knowing the supplier’s cost and the retailer’s benefit of effort and the distribution of the retailer’s quantity needs. Firms would need to share such information to construct a contract that sustains efficient effort while minimizing payment adjustment costs. If each firm is concerned that sharing such information weakens its bargaining position, it would be interesting to study conditions in which channel partners can overcome this challenge to formulate an efficient contract (see Wernerfelt 2012).

Our research implies that partners should cooperate to formulate a relationship that maximizes joint surplus while avoiding situations in which one party has an incentive to violate the terms of the contract or in which the parties need to make costly adjustments to those terms. An important managerial implication of our model is that a supplier offering a full product line, including basic goods that do not require customization, can increase profits for both firms in a supply relationship. Trading multiple goods allows more flexible payment arrangements, which allow firms to sustain greater effort and higher total premium payments than would be possible if the supplier provided only a customizable good. The additional surplus generated when firms trade both customizable and basic goods implies that both firms in the relationship can earn higher profits.

Acknowledgments
Helpful comments were provided by Sharmila Chatterjee, Bob Gibbons, Brandon Park, Sarah Park, Catherine Tucker, Birger Wernerfelt, Juanjuan Zhang, and seminar participants at the University of California, Berkeley; Chapman University; University of Chicago; Emory University; Georgia Institute of Technology; Massachusetts Institute of Technology; University of North Carolina; University of Toronto, the Marketing Science Conference, and the University of Texas Dallas FORMS conference. The authors thank the restaurant supply company that provided the price information that motivated this paper.

Endnotes
1 This example is based on a report by the International Foodservice Distributors Association (Caldwell 2017).

2 As stated by the company’s chief executive officer, “You would expect that the customer pays us high prices for the value-added services we provide. Well, that doesn’t happen. . . . We cross-sell our other products to these customers by offering them significant breaks on the value-add products in return for their commitment to buy the book-and-ship products exclusively from us. In a way, in these relationships the commodity products subsidize the specialty products” (Narayandas 1998, p. 8).

3 See Proposition 1 in Section 3.2 for additional detail.

4 For example, if there are two products and two possible demand states, then placing a large price premium on a product with volatile demand implies a large difference in the total premium payment between the high- and low-demand states. By contrast, placing a smaller (or even negative) price premium on this volatile product, with a larger premium on a more stable product, leads to a small difference in the premium payment between the two demand states (see the numerical example in Section 3.3 for more detail).

5 Without loss of generality, we have set marginal production costs equal to zero, so the only costs in the model are effort costs.

6 The seller’s utility represents the profits generated from the units of the goods it otherwise would have sold to the retailer if they had traded with each other instead of on the outside market.

7 Recall that for simplicity, we assume that outside market prices are constant over time. If we instead allowed outside market prices to vary over time, then results equivalent to ours would hold if we allowed contracts with a constant per-unit price premium for each product, that is, a constant price markup over market prices.

8 Most of the theoretical literature on relational contracting assumes that the agent decides effort before the principal decides the bonus payment, in which case the principal can benefit from the effort but then shirk on the bonus (e.g., Levin 2003). Using this alternative timing assumption would not significantly affect our key results. We believe that the timing assumption used in this paper, in which the payment is decided before effort, is more realistic in the context of supply chain relationships.

9 There is a bad equilibrium in which the buyer always buys from the outside market and the supplier always exerts zero effort.

10 Our model also allows for the possibility of other equilibria in which firms trade on the outside market during some periods, for example, when the total premium payment specified in the contract exceeds the buyer’s willingness to pay. Note that trading on the outside market does not constitute punishment if it occurs on the equilibrium path.

11 In the model extension that follows, there are four demand states and two prices, so the firms sometimes need to incur an adjustment cost to achieve their desired payment level.

12 If the benefit of effort is very large (β > 3.5), then condition (15) does not hold for any value of Hf. In this case, firms can always sustain optimal effort, and it is optimal to place a price premium only on the customizable good even as Hf grows without bound.

13 The assumption ω < 0.5 has no effect on the results in this section, but larger values of ω could change the optimal contract with order splitting derived in the next section.

14 By contrast, if the correlation in product demands is weak enough that (31) is violated, there is an efficient contract that places a price premium only on the customizable good and involves payment adjustments either when demand for the customizable good is high or when it is low, similar to the contract derived in the next section.

15 By contrast, if (24) holds, the optimal contract from previous sections involves a positive price premium for the basic good but a negative price premium for the customizable good. In this case, firms
could use a contract like the one derived in the preceding section because order splitting would allow both upward and downward adjustments to the total payment given that the buyer could split its order for either good.

16 Because $\phi$ is a function of $F_{t_{max}}$, there is not a simple closed-form expression that solves (43) for $F_{t_{max}}$. If there are multiple solutions, then the greatest value of $F_{t_{max}}$ that solves (43) generates the greatest surplus and is therefore efficient.

17 Alternatively, if demand for products were not perfectly correlated, firms could use the fixed payment and per-unit prices to match the optimal payment during some states while trading on the outside market or accepting suboptimal payment levels during other states.

References


