Facilitating Demand Risk-Sharing with the Percent Deviation Contract

Matthew J. Drake
A.J. Palumbo School of Business Administration
John F. Donahue Graduate School of Business
Duquesne University
Pittsburgh, PA 15282, USA
drake987@duq.edu

Julie L. Swann
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0205, USA
jswann@isye.gatech.edu

June 2007

Facilitating Demand Risk-Sharing with the Percent Deviation Contract

Suppliers do not have much incentive to build capacity for supply chains in which the buyer bears little or no demand risk. This hinders the supply chain from satisfying the optimal amount of customer demand from a channel perspective. We proposed the percent deviation contract as a mechanism to improve the overall performance of this type of supply chain. This contract induces a dynamic game of perfect information, and we characterize the subgame-perfect Nash Equilibria under various contract scenarios. We establish ways to set the contract parameters to coordinate the supply chain and show that the percent deviation contract is able to achieve channel coordination in some cases where the quantity flexibility contract fails. In order to aid the implementation of the percent deviation contract in practice, we develop ways to set the parameters to satisfy the buyer's individual-rationality constraint. We conclude with a numerical example based on industry data that highlights the important results.

1. Introduction

The proliferation of computerized information systems in the 1990s facilitated the establishment of supply chain partnerships in which demand information is shared between firms. The upstream firms can use this information to reduce the traditional demand distortion due to the bullwhip effect (Lee et al., 1997). Some firms have also incorporated this information into contracts that induce their supply chain partners to share demand risk, thereby improving supply chain efficiency. Many researchers and practitioners (e.g., Lee (2004) and Finley and Srikanth (2005)) have advocated demand risk-sharing as a necessary condition for supply chain collaboration efforts to be successful in practice. In this paper we analyze one such contracting mechanism, which we denote as the *percent deviation* contract.

The percent deviation contract is most applicable to supply chains in which the buyer does not traditionally have any demand risk since she places orders with the supplier only when her demand is known with certainty; that is, the buyer does not carry any excess inventory. This environment often occurs in many service operations. One industry that stands to benefit from application of the percent deviation contract is truckload transportation. In fact, a large truckload carrier originally proposed the idea for this contract structure but did not know how to set the parameters or whether or not the contract would be beneficial. While these carriers generally have standing weekly orders for loads with their bigger customers, many shippers call dispatch requesting a pickup in a few hours. This limits the carrier's ability to utilize his equipment effectively by allocating trailers in advance or coordinating backhauls from prior shipments. The percent deviation mechanism is applicable in traditional manufacturer-retailer channels as well, such as home construction, equipment integrators, window replacement, or door-to-door sales. In all of these industries, the supplier bears most, if not all, of the demand risk in many arrangements.

We analyze the strategic properties of the percent deviation contract in which the buyer gives an initial order estimate and the supplier pre-acquires inventory at a low cost. Once the buyer's customer demand is realized, she places her final order, and the supplier fulfills all or a portion of the order, possibly by expediting, at a higher cost. The buyer pays a penalty if her final order is outside of an allowable range established around her initial order estimate. We characterize the subgame-perfect Nash Equilibria decisions when the supplier has a fixed expediting capacity and discuss methods of channel coordination to optimize the performance of the entire system. Since the buyer assumes some demand risk under the percent deviation contract, her expected profit may be less than that under a traditional contracting structure; therefore, we develop a method that the supplier can use to satisfy the buyer's individual-rationality constraint.

Our contribution to the existing literature on supply chain collaboration includes analysis of a risk-sharing contract where decisions made by the buyer and supplier explicitly depend on each other and are solvable in the framework of a dynamic, extensive form game. This dynamic game necessarily results in a more complex contract, but we also show that this contract can be strictly Pareto-improving for both parties. Our contract has a structure similar to the quantity flexibility contract, but ours does not enforce limits on the buyer's final behavior, so we show that this contract can coordinate the supply chain in some cases where quantity flexibility cannot. Many contracting models consider a supply chain that has infinite capacity; whereas, the supply chain capacity in our model is a function of the supplier's decisions as well as an external constraint.

2. Literature Review

The breadth of supply chain contracting literature has grown significantly over the last two decades as researchers and practitioners have examined strategic relationships between supply chain partners. (See Tsay et al. (1998) for a review of traditional contracting mechanisms.) One stream of supply chain contracting literature has proposed and analyzed methods of coordinating decentralized decisions to attain the optimal supply chain profit. Examples of these studies include Weng (1997), Parlar and Weng (1997), Taylor (2002), and Huggins and Olsen (2003). We discuss below the most relevant contracting references, which model a system with multiple, sequential decisions.

Tsay (1999) analyzes a quantity flexibility contract in which the retailer commits to purchasing no less than a certain percentage of the initial forecast while the supplier agrees to fulfill up to a certain percentage above the forecast. He also evaluates the sharing of demand risk that produces the coordinated channel. Tsay and Lovejoy (1999) extend these results to a rolling horizon decision environment. In contrast to quantity flexibility, the percent deviation contract places no limits on the buyer's final order, although it adds complexity to the decision environment by including additional contract parameters. We show in Section 4.3 that this added complexity can be justified because the percent deviation contract succeeds in coordinating the supply chain in several cases where the quantity flexibility contract is known to be unable to coordinate the channel.

Donohue (2000) and Cachon (2004) analyze contracts with two-tier pricing structure that induce early commitment from buyers. In both of these contracts the buyer is bound to her order in both periods, whereas in our contract the first order is only an estimate of demand and can be freely adjusted once demand is known. These two papers only consider the full compliance contract regime where the supplier must fulfill the entire order; whereas, we model the supplier's compliance decision explicitly.

Several contracts employ an options framework where the buyer makes a firm order commitment and purchases options for additional goods to be exercised if demand is high. Cachon and Lariviere (2001) consider a single period model with options and forecast sharing. Since the buyer has an incentive to provide a biased forecast, they develop conditions that facilitate the credible sharing of forecasts under both full and voluntary compliance. Barnes-Schuster et al. (2002) extend the options framework using a two-period model with correlated demand between periods. In our model we have no firm commitment and no upper bound on the final order amount, so our contract cannot be reduced to an option-based model.

A recent series of studies (see, for example, Jin and Wu (2001) and Erkoc and Wu (2005)) have analyzed reservation fee supply contracts in which the buyer pays a (usually) deductible fee to reserve capacity along with an exercise fee for the final order quantity. The manufacturer builds capacity based on the reservations made, but they can also build excess capacity to offer at a higher spot rate once demand is realized. The aforementioned studies only consider *linear* reservation fee contracts—where each unit ordered is charged the same prices. The percent deviation contract is a special case of a *piecewise-linear* reservation fee contract in which the reservation and exercise prices differ for various portions of the order.

In addition to contracting, several papers (e.g. Lee et al. (2000), Cachon and Fisher (2000), and Balakrishnan et al. (2004)) have examined various ways of reducing the bull-whip effect through information sharing in decentralized supply chains. Kulp et al. (2004) study the benefits the manufacturer gains under different degrees of information sharing and collaboration. They find that most of the manufacturers' benefit from information integration comes from collaborative activities such as vendor-managed inventory and collaborative forecasting instead of simply sharing information. On the contrary, our results suggest that the risk sharing induced by the percent deviation contract enables the supplier and the entire channel to attain higher profit.

The next section develops the model for the general case where the supplier has an expediting capacity constraint as well as the special case of infinite capacity. Section 4 develops conditions on the contract parameters that satisfy each party's participation constraints, details ways to coordinate the channel in each decentralized scenario, and compares the percent deviation contract with the well-known quantity flexibility contract. Section 5 provides the results of a numerical example using a demand distribution estimated from a major manufacturer's weekly shipping activity. Conclusions and suggestions of future research are given in Section 6.

3. Models and Scenarios

The percent deviation contract accommodates the following sequence of decisions. The buyer provides an initial estimate of its final-order demand that will be placed at a later date. The seller can then use this information to acquire goods in advance (e.g., a truckload carrier can preposition trucks or coordinate backhauls to optimize his transportation network) at a low cost in anticipation of this demand. When the buyer's demand is known with certainty, she places her actual order with the supplier. Depending on the contract parameters, the seller can choose to satisfy additional demand by expediting or subcontracting at a high cost or can choose to fulfill only the demand equal to the number of previously-acquired goods. The percent deviation penalty is the mechanism that punishes the buyer for unrealistic estimates. If the buyer's final order is within a certain percentage above or below her initial estimate, no penalty is charged. If the order exceeds the limits, the supplier charges a penalty on all goods ordered outside of the tolerable range.

3.1 Notation and Assumptions

We employ notation adapted from Donohue (2000). The buyer receives r dollars in revenue for each unit, and she pays a wholesale price, w, to the supplier. We assume that the buyer earns a positive gross margin from these transactions (i.e., $r > w$). Consumer demand for a period is given by the random variable X, which has a continuous, differentiable probability distribution function, $f(x)$. If the buyer cannot satisfy her customers' demand (due to lack of product availability), she incurs a customer penalty of β per unit. This β could also be viewed as the higher cost from using an alternative supplier not under long-term contract.

The seller faces a cost of c_1 dollars to acquire goods in anticipation of demand and must pay c_2 dollars to satisfy demand after the firm order has been placed. We assume that $c_2 > c_1$, so the c_2 can be thought of as an expediting or subcontracting cost. If the supplier has excess inventory at the end of the period, he receives a unit salvage value of v. It is natural to assume that $w > v$ and $c_1 > v$, which ensure that the supplier does not receive too much of a benefit from selling goods for salvage. Since the seller may choose not to satisfy the buyer's entire order, he must pay the buyer α for each unit ordered but not delivered.

We assume that $\alpha < \beta$, which signifies that lost customers are more costly for the buyer than for the supplier.

The per-unit penalty that the buyer must pay the supplier for orders outside of the allowable deviation range is denoted by p, while $d \in [0, 1]$ is the percentage that defines the range. For orders above the upper limit of the range, the buyer only pays for the units that the supplier fulfills. The buyer's initial forecast of her order is given by q_1, and the actual order is q_2. The number of units the supplier acquires in advance of demand is t_1, and the additional goods expedited or subcontracted are denoted by t_2. The supplier has a maximum expediting capacity of M units.

This particular way of modeling the supplier's capacity bears further consideration. It is important to note that the capacity for the supplier's pre-acquisition decision is infinite. By setting the t_1 value, the supplier is *de facto* determining the capacity of the system as a whole, which is equal to $t_1 + M$. This structure is appropriate for buyer-supplier transactions in which the supplier has a lot of capacity in his system, but he must make the allocation decisions over many customers before production occurs. Therefore, if the supplier knows that a particular buyer will require many units in a given period, he can allocate sufficient capacity to satisfy the large order; closer to the purchase date, however, he can only provide a limited amount of excess capacity if necessary because the rest of his system is dedicated to fulfilling orders from other customers.

We make the following assumptions that improve tractability but are not likely to impede the application of the results. The first assumption is that all costs are linear per unit of demand for a single product line because we are interested in the structure of the incentives. Another assumption is the existence of complete, symmetric cost, capacity, and demand information between the two parties. (We relax this assumption in Drake (2006) by considering a buyer that has a private demand forecast following the structure developed by Ferguson et al. (2005).) When the buyer places her final order, she knows the exact demand as is usual in the other relevant channels discussed in Section 1.

If the actual demand amount exceeds the upper limit of the deviation range, the cost and penalty parameters determine whether or not the buyer's order equals the full demand. In order for the buyer to order above the deviation threshold, the net cash flow from satisfying the demand must exceed the penalty that she must pay her customer for not satisfying demand. If the inequality

$$r - w - p > -\beta \tag{1}$$

holds, then $q_2 = X$, or the actual customer demand. If this inequality is not satisfied (for instance, if the penalty for ordering outside of the deviation range is too high), $q_2 = min\{X, (1 + d)q_1\}$. The additional assumption $w > p$ assures that if the actual demand is below the lower limit of the deviation range, $(1 - d)q_1$, the buyer orders the actual demand.

3.2 General Model with Finite Expediting Capacity

We begin our analysis with the decentralized structure in which each party makes decisions to optimize his individual expected profit. Even though the supplier has an expediting capacity of M units, the cost of expediting these units, c_2, might be too high for him to choose to do so. In order for the supplier to use any of this expediting capacity, the cash flow from expediting must be higher than the cost of failing to expedite. These flows are dependent on whether or not the supplier will receive the deviation penalty on some or all of these units. While the two buyer scenarios discussed above, which are dependent on whether or not the the buyer is willing to place orders above the upper limit of the deviation range, generate different supplier responses, the derivation and form of these optimal decisions are the same in both cases. Thus, in this paper we only examine the case in which the buyer is willing to order the entire demand even if she must pay the deviation penalty. (See Drake (2006) for the case in which the buyer will not order above the upper limit of the deviation range.)

When the buyer is willing to order the actual demand, the supplier's expediting decision can be determined *a priori*, without knowledge of how many units for which the buyer will pay the deviation penalty. If $w - c_2 > -\alpha$, then the supplier finds it beneficial to expedite whether or not he receives the deviation penalty on any units; consequently, $t_2^* = (\min\{q_2^* - t_1, M\})^+$. We will denote this as Case A. Similarly, if $w - c_2 + p < -\alpha$, the supplier would not choose to expedite any units even if he were receiving the deviation penalty on all of the units; and thus, $t_2^* = 0$. This will be Case B.

We will use backward induction to solve for the subgame-perfect Nash Equilibria in each scenario. We now formulate the expected profit functions for the buyer and supplier in Case

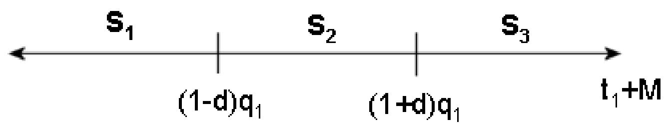


Figure 1: Regions of capacity defining the form of the supplier's expected profit function

A, where $q_2^* = X$ and $t_2^* = (\min\{q_2^* - t_1, 0\})^+$. The supplier chooses t_1 to maximize his expected profit function:

$$\begin{aligned}
\Pi_A^S &= w\left[\int_0^{t_1+M} xf(x)dx + (t_1+M)\left(1-F(t_1+M)\right)\right] + \\
&\quad p\int_0^{\min\{t_1+M,(1-d)q_1\}} \left(\min\{t_1+M,(1-d)q_1\} - x\right) f(x)dx + \\
&\quad p\left[\int_{(1+d)q_1}^{t_1+M} (x-(1+d)q_1)f(x)dx + \left(t_1+M-(1+d)q_1\right)^+ \left(1-F(t_1+M)\right)\right] + \\
&\quad v\int_0^{t_1}(t_1-x)f(x)dx - c_1t_1 - c_2\left[\int_{t_1}^{t_1+M}(x-t_1)f(x)dx + M\left(1-F(t_1+M)\right)\right] - \\
&\quad \alpha\int_{t_1+M}^{\infty}(x-t_1-M)f(x)dx. \qquad (2)
\end{aligned}$$

There are three separate functions that represent realizations of the expected profit function in (2) based on the relationship between the total system capacity, $t_1 + M$, and the boundaries of the deviation range, $(1-d)q_1$ and $(1+d)q_1$. The expected profit regions are depicted in Figure 1. In region S_1 the supplier sets capacity so that he cannot even satisfy the lower limit of the deviation range. Region S_2 prescribes that the total system capacity lies somewhere in the deviation range. In these first two regions, the buyer will never pay the deviation penalty for orders above the upper limit of the range because these units will never be fulfilled. Region S_3 specifies that the system capacity exceeds the upper limit of the deviation range. In each region only one of the three separate expected profit realizations is feasible, regardless of which expected profit is higher in the region. The following observation establishes that the overall expected profit function is continuous.

Observation 1 *The individual expected profit function realizations that are active in two adjacent feasible regions are equal at the boundary (i.e.,* $\Pi^S_{A.I.}((1-d)q_1 - M) = \Pi^S_{A.II.}((1-d)q_1 - M)$ *and* $\Pi^S_{A.II.}((1+d)q_1 - M) = \Pi^S_{A.III.}((1+d)q_1 - M)$*).*

Lemma 1 *If $w + \alpha > p + c_2$, the supplier's expected profit function in (2) is piecewise-concave, a continuous, piecewise function whose separate segments are individually concave.*

Each of the individual realizations of (2) has a corresponding maximizing value that is derived from the solution to the following equations,

$$t_1^{A.I.} \in \left\{ t : F(t+M) = \frac{w+\alpha-c_1-(c_2-v)F(t)}{w+\alpha-c_2-p} \right\} \tag{3}$$

$$t_1^{A.II.} \in \left\{ t : F(t+M) = \frac{w+\alpha-c_1-(c_2-v)F(t)}{w+\alpha-c_2} \right\} \tag{4}$$

$$t_1^{A.III.} \in \left\{ t : F(t+M) = \frac{w+\alpha-c_1+p-(c_2-v)F(t)}{w+\alpha-c_2+p} \right\}, \tag{5}$$

which are all dependent on the supplier's available expediting capacity. It is interesting to note that all of these values are *independent* of the buyer's order quantity; they can easily be computed from the exogenous parameters. We can always solve these equations by applying the Intermediate Value Theorem since the left-hand sides are all bounded between 0 and 1.

To understand the supplier's best response, we can consider the individual functional maximizers in (3)–(5) and their relationship to each other and the boundaries of the feasible regions. The following theorem characterizes the supplier's best response, which is dependent on the specific value of the buyer's decision, q_1, via the feasible region boundary conditions. Figure 3 in the Appendix provides a graphical tree of decisions for each specific set of parameter values. Proofs of major results are in the Appendix; additional proofs are available upon request.

Theorem 1 *The supplier's best response to a given value of q_1 when $w + \alpha > p + c_2$ is*

$$t_1^*(q_1) = \begin{cases} t_1^{A.I.}, & \text{if } t_1^{A.I.} \leq (1-d)q_1 - M \text{ \& } t_1^{A.III.} \leq (1+d)q_1 - M; \\ t_1^{A.II.}, & \text{if } (1-d)q_1 - M \leq t_1^{A.II.} \leq (1+d)q_1 - M \text{ \& } \\ & t_1^{A.III.} \leq (1+d)q_1 - M; \\ t_1^{A.III.}, & \text{if } t_1^{A.II.} \geq (1+d)q_1 - M \text{ \& } t_1^{A.III.} \geq (1+d)q_1 - M; \\ (1-d)q_1 - M, & \text{if } t_1^{A.II.} \leq (1-d)q_1 - M \leq t_1^{A.I.} \text{ \& } \\ & t_1^{A.III.} \leq (1+d)q_1 - M; \\ \arg\max_{t_1^{A.I}, t_1^{A.III.}} \Pi_A^S, & \text{if } t_1^{A.I.} \leq (1-d)q_1 - M \text{ \& } t_1^{A.III.} \geq (1+d)q_1 - M; \\ \arg\max_{(1-d)q_1 - M, t_1^{A.III.}} \Pi_A^S, & \text{if } t_1^{A.II.} \leq (1-d)q_1 - M \leq t_1^{A.I.} \text{ \& } \\ & t_1^{A.III.} \geq (1+d)q_1 - M; \\ \arg\max_{t_1^{A.II.}, t_1^{A.III.}} \Pi_A^S, & \text{if } (1-d)q_1 - M \leq t_1^{A.II.} \leq (1+d)q_1 - M \leq t_1^{A.III.}. \end{cases} \tag{6}$$

Table 1: Possible SPNE pairs and feasibility conditions for Case A's explicit $t_1(q_1)$ decisions

$(\mathbf{q_1^*}, \mathbf{t_1^*}(\mathbf{q_1^*}))$	**Feasibility Conditions**
$\left(q_1^{A.I.} \equiv \left\{q : q \geq \max\left\{\frac{t_1^{A.I.}+M}{1-d}, \frac{t_1^{A.III.}+M}{1+d}\right\}\right\}, t_1^{A.I.}\right)$	Always Feasible
$\left(q_1^{A.II.} \equiv \max\left\{\frac{F^{-1}(0)}{1-d}, \frac{\max\left\{t_1^{A.II.}, t_1^{A.III.}\right\}+M}{1+d}\right\}, t_1^{A.II.}\right)$	Always Feasible
$(q_1^{A.III.} \equiv \{q : (1-d)F((1-d)q) = (1+d)(1-F((1+d)q))\}, t_1^{A.III.})$	$q_1^{A.III.} \leq \frac{\min\left\{t_1^{A.II.}, t_1^{A.III.}\right\}+M}{1+d}$
$(q_1^{I.A.IV.} \equiv \left\{q : F((1-d)q) = \frac{r-w-\alpha+\beta}{r-w-\alpha+\beta+p}\right\}, (1-d)q_1^{A.IV.} - M)$	$\max\left\{\frac{t_1^{A.II.}+M}{1-d}, \frac{t_1^{A.III.}+M}{1+d}\right\} \leq q_1^{A.IV.}$ $q_1^{A.IV.} \leq \frac{t_1^{A.I.}+M}{1-d}$

The buyer must choose the q_1 that maximizes her expected profit while anticipating the supplier's response to her chosen value. The buyer's expected profit function is given by

$$\begin{aligned}
\Pi_A^B \quad = \quad & (r-w)\left[\int_0^{t_1^*(q_1)+M} xf(x)dx + (t_1^*(q_1)+M)\left(1-F\left(t_1^*(q_1)+M\right)\right)\right] - \\
& p\int_0^{\min\{(1-d)q_1, t_1^*(q_1)+M\}} \left(\min\{(1-d)q_1, t_1^*(q_1)+M\} - x\right) f(x)dx - \\
& p\left[\int_{(1+d)q_1}^{t_1^*(q_1)+M} (x-(1+d)q_1)f(x)dx + \left(t_1^*(q_1)+M-(1+d)q_1\right)^+ \left(1-F\left(t_1^*(q_1)+M\right)\right)\right] + \\
& (\alpha-\beta)\int_{t_1^*(q_1)-M}^{\infty} \left(x - t_1^*(q_1) - M\right) f(x)dx. \qquad (7)
\end{aligned}$$

The supplier's best response function in (6) is comprised of four explicit values as well as three situations where the supplier chooses the profit-maximizing quantity from a set of two of the explicit values. We first determine how the buyer should set q_1 when the supplier will respond with each of the four possible t_1 values. Each of these cases results in a different realization of the buyer's expected profit function in (7). We apply the Karush-Kuhn-Tucker (KKT) conditions (c.f. Bazaraa et al. (1993: 151–55))) over each realization's feasible range of decisions to determine the constrained optimal values of q_1.

Theorem 2 *The possible subgame-perfect Nash Equilibrium (SPNE) decision pairs for explicit $t_1(q_1)$ decisions are given in Table 1.*

We must also determine the buyer's optimal decision over the regions where she knows that the supplier will be selecting the maximizing argument from a set of two values. From

Theorem 1, we establish the following ranges of q_1 under which each of these situations is possible.

$$\arg\max_{t_1^{A.I},t_1^{A.III.}} \Pi_A^S : \quad \frac{t_1^{A.I.}+M}{1-d} \le q_1 \le \frac{t_1^{A.III.}+M}{1+d} \tag{8}$$

$$\arg\max_{(1-d)q_1-M,t_1^{A.III.}} \Pi_{A.}^S : \quad \frac{t_1^{A.II.}+M}{1-d} \le q_1 \le \min\left\{\frac{t_1^{A.I.}+M}{1-d}, \frac{t_1^{A.III.}+M}{1+d}\right\} \tag{9}$$

$$\arg\max_{t_1^{A.II.},t_1^{A.III.}} \Pi_A^S : \quad q_1 \le \frac{t_1^{A.II.}+M}{1-d} \;\&\; \frac{t_1^{A.II.}+M}{1+d} \le q_1 \le \frac{t_1^{A.III.}+M}{1+d} \tag{10}$$

Since all three situations involve the possible decision $t_1^{A.III.}$, we define the *difference function*, $\Delta(t) \equiv \Pi_A^S(t_1^{A.III.}) - \Pi_A^S(t)$, where t is any other possible supplier pre-acquisition amount. While it is difficult to determine the exact feasible region for q_1 that induces each of the possible $t_1(q_1)$ values, we can use these difference functions to explain how a buyer would determine her optimal decision for a given set of parameters.

Proposition 1 *The structure of the three difference functions for the decisions in (8)–(10) enables us to determine the specific ranges of q_1 that induce each of the two possible supplier values for t_1.*

For each instance there are at most seven decision pairs from Table 1 and obtained from the procedure in Proposition 1, but some of these decisions may not be mutually feasible given a set of problem parameters. Since this set contains a maximum of seven elements, the buyer can evaluate her expected profit in (7) with respect to each of the feasible pairs and select the initial order estimate that yields the highest expected profit to obtain the overall subgame-perfect Nash Equilibrium for this sequential supply chain game.

Each of the formulas used in computing the potential decision pairs has an economic interpretation. The buyer always sets her initial order estimate with a goal of minimizing the expected deviation penalty that she pays under each scenario. In the case where she can conceivably experience a lower and an upper deviation penalty, the quantity she chooses balances the expected lower and upper deviation penalties. The supplier also seeks to balance the expected revenue that he receives from pre-acquiring inventory with the cost of doing so as well as the expected expediting and shortage costs. Even though the resulting formulas are more complicated, the supplier follows the same rationale as in a traditional newsvendor contract.

We now consider Case B, where the supplier chooses not to expedite any units after the buyer places her final order because of a high expediting cost. This case is analogous to Case A when $M = 0$ since the supplier can be viewed as having an effective expediting capacity of zero units if he chooses not to expedite. The maximizing values below correspond with the equations in (3)–(5) with $M = 0$.

$$t_1^{B.I.} \in \left\{ t : F(t) = \frac{w + \alpha - c_1}{w + \alpha - v - p} \right\} \tag{11}$$

$$t_1^{B.II.} \in \left\{ t : F(t) = \frac{w + \alpha - c_1}{w + \alpha - v} \right\} \tag{12}$$

$$t_1^{B.III.} \in \left\{ t : F(t) = \frac{w + \alpha - c_1 + p}{w + \alpha - v + p} \right\} \tag{13}$$

Using the relationships in Lemma 6 (stated in the Appendix) to simplify the feasibility conditions, the supplier's best response in this scenario is characterized by the following theorem. While it may not seem like it at first glance, the feasibility conditions for each of the decisions in (14) correspond to those in (6).

Theorem 3 *The supplier's best response to a given value of q_1 when $w + \alpha > p + v$ is*

$$t_1^*(q_1) = \begin{cases} t_1^{B.I.}, & \text{if } t_1^{B.I.} \leq (1-d)q_1 \ \&\ t_1^{B.III.} \leq (1+d)q_1; \\ t_1^{B.II.}, & \text{if } (1-d)q_1 \leq t_1^{B.II.} \ \&\ t_1^{B.III.} \leq (1+d)q_1; \\ t_1^{B.III.}, & \text{if } t_1^{B.II.} \geq (1+d)q_1; \\ (1-d)q_1, & \text{if } t_1^{B.II.} \leq (1-d)q_1 \leq t_1^{B.I.} \ \&\ t_1^{B.III.} \leq (1+d)q_1; \\ \arg\max_{t_1^{B.I.}, t_1^{B.III.}} \Pi_B^S, & \text{if } t_1^{B.I.} \leq (1-d)q_1 \ \&\ t_1^{B.III.} \geq (1+d)q_1; \\ \arg\max_{(1-d)q_1, t_1^{B.III.}} \Pi_B^S, & \text{if } t_1^{B.II.} \leq (1-d)q_1 \leq t_1^{B.I.} \ \&\ \\ & t_1^{B.III.} \geq (1+d)q_1; \\ \arg\max_{t_1^{B.II.}, t_1^{B.III.}} \Pi_B^S, & \text{if } (1-d)q_1 \leq t_1^{B.II.} \leq (1+d)q_1 \leq t_1^{B.III.}. \end{cases} \tag{14}$$

The possible subgame-perfect Nash Equilibrium decision pairs in case I.B. are the same as those for Case A (given in Table 1) except with the B supplier decision values replacing the A decisions. The buyer's decisions are exactly the same as those in Case A since the actual q_1 values are independent of M. The buyer can again apply the methods described in the proof of Proposition 1 to determine her optimal initial order estimate in the cases in which the supplier's best response is the value among a set of maximizing arguments for two of the expected profit realizations. Once the feasible set of possible decision pairs is determined, the buyer can again substitute each of them into her expected profit function to find the maximizing decision pair, which is the SPNE.

If the contract's parameters are such that neither of the above scenarios (A or B) apply, then the supplier's expediting decision is dependent on the magnitude of the final order. In order for it to be Pareto optimal for the supplier to fulfill the order, the cash flow from satisfying must be greater than the cash flow from not satisfying, or $(w-c_2)(q_2-t_1)+p(q_2-(1+d)q_1) > -\alpha(q_2-t_1)$. Solving for q_2, the seller satisfies the extra demand if

$$q_2 > \frac{(w-c_2+\alpha)t_1 + p(1+d)q_1}{w-c_2+\alpha+p} \equiv \Lambda.$$

Formally, the supplier's expediting decision is

$$t_2^* = \begin{cases} \left(\min\{q_2 - t_1, M\}\right)^+, & \text{if } q_2 > \Lambda \\ 0, & \text{if } q_2 < \Lambda. \end{cases}$$

The dependence of supplier compliance on the magnitude of the final order may be problematic for both parties. The buyer must wait until the demand realization to know if the supplier is going to comply with the entire order, causing her added supply uncertainty. Knowing that the expediting decision rests with the magnitude of the order might induce the buyer to inflate her final order so that the supplier will fulfill the entire amount. The buyer's strategic behavior is detrimental for the supplier because he could be induced to expedite when he would not otherwise.

To alleviate these difficulties, we recommend that the parties set the negotiated parameters—p and α—such that the contract assumes another case. This could be accomplished by letting $\alpha > c_2 - w$, shifting the contract to the A case. Of course, shifts to other scenarios are possible through negotiations, depending on the relative market power of the parties. Since both parties have an incentive to set the contract parameters to move the contract to other cases, we omit this situation from the analysis.

3.3 Infinite Expediting Capacity

We conclude our presentation of the general models by considering a special case in which the supplier's expediting capacity is infinite (or especially large for practical purposes). These uncapacitated models have an especially simple structure that enables us to develop (quasi) closed-form optimal decisions. Since this case is based on the expediting capacity, we must only develop models for case analogous to A above, in which the supplier chooses to expedite. It does not matter how much extra capacity the supplier has if he chooses not to use it.

Again we only consider the case where the buyer orders the actual demand in all cases (i.e., $q_2^* = X$ and $t_2^* = (\min\{q_2 - t_1, M\})^+$) since (1) holds and $w - c_2 > -\alpha$. (It is straightforward to extend these models to the case in which the buyer does not order above the deviation range; see Drake (2006) for details.) We denote this scenario as A.∞. The supplier's expected profit function is again the same as in (2) with $M = \infty$, but this substitution results in the simpler function

$$\begin{aligned}\Pi^S_{A.\infty} &= w\int_0^\infty xf(x)dx + p\int_0^{(1-d)q_1}((1-d)q_1 - x)f(x)dx + p\int_{(1+d)q_1}^\infty (x-(1+d)q_1)f(x)dx + \\ &\quad v\int_0^{t_1}(t_1 - x)f(x)dx - c_1t_1 - c_2\int_{t_1}^\infty (x-t_1)f(x)dx. \qquad (15)\end{aligned}$$

The expected profit function in (15) is concave by Lemma 1, so we can solve for the optimal pre-acquisition amount using first order conditions. This yields

$$t_1^{A.\infty} \in \left\{t : F(t) = \frac{c_2 - c_1}{c_2 - v}\right\}, \qquad (16)$$

which is independent of the buyer's initial order estimate because of the symmetric information assumption and the infinite total system capacity.

Similarly, the buyer's expected profit function is the same as in (7), but infinite supplier capacity yields the simplified form.

$$\Pi^B_{A.\infty} = (r-w)\int_0^\infty xf(x)dx - p\int_0^{(1-d)q_1}((1-d)q_1 - x)f(x)dx - p\int_{(1+d)q_1}^\infty (x-(1+d)q_1)f(x)dx \qquad (17)$$

Because the supplier is willing to expedite to satisfy the buyer's order regardless of its size, the buyer's expected profit function is no longer dependent on the supplier's t_1 decision. In this case, the t_1 decision only affects the supplier's profitability and not her ability to fulfill the buyer's order.

The buyer's optimal initial order estimate is given by $q_1^{A.\infty} \equiv \{q : (1-d)F((1-d)q) = (1+d)(1-F((1+d)q))\}$, which corresponds with her optimal decision in case A.III. in which the supplier is able to satisfy orders above the upper limit of the deviation range. The decision $q_1^{A.\infty}$ is the value of q_1 that equates the marginal expected deviation penalty for demands below the lower limit of the range, $p(1-d)F((1-d)q_1)$, to the marginal expected penalty for orders above the upper limit, $p(1+d)(1-F((1+d)q_1))$. Since the nominal deviation penalty, p, is the same regardless of whether the deviation was a lower deviation or an upper deviation, it is irrelevant to the buyer's decision. Of course, if there were two deviation penalties, p_l and p_u, they would affect the buyer's decision.

Thus, the SPNE for the A.∞ case is $(q_1^*, t_1^*(q_1^*)) = \left(q_1^{A.\infty}, t_1^{A.\infty}\right)$ for all parameter sets such that $w - c_2 > -\alpha$. It applies in situations where the supplier always has enough extra capacity in her network to satisfy the buyer's order. It would be most reasonable when the buyer's requirements are small compared with the supplier's capabilities. Consequently, the supplier would only need to utilize the more complicated capacitated contracts for customers who require a large portion of his capacity. Since these buyers are larger, they are presumably more important to the supplier, so he would have more incentive to utilize a more complicated contract for these customers.

4. Economic Analysis and Model Extensions

4.1 Individual Rationality Constraints

The practical implementation of the percent deviation contract is necessarily impacted by the competitive power of the parties. If the buyer has a powerful market presence, she will likely be able to negotiate favorable contract terms by threatening to use another supplier who offers a more traditional agreement. (We assume that the contract is used in a competitive industry, so the buyer can find another supplier with comparable service performance and quality.) The terms of the contract, therefore, must satisfy the buyer's individual-rationality constraint, which says that under the percent deviation contract she must be able to attain an expected profit at least as great as she could under a traditional mechanism. See Tirole (1988) for a detailed discussion of individual-rationality constraints. If this constraint is not satisfied, she will switch to another supplier.

In this section we compare our percent deviation contract to the *status quo* of a traditional wholesale-price contract. In the cases where the supplier's expediting capacity is limited, the percent deviation mechanism can induce the supplier to pre-acquire significantly more inventory than he would under the wholesale-price contract. This additional ability to meet demand is beneficial for both parties, resulting in higher expected profits without further contract modifications. In situations where the supplier does not increase his pre-acquisition quantity significantly (i.e., the deviation penalty is not high enough to induce him to pre-acquire much more inventory), it is clear that the buyer will earn less expected profit under the percent deviation contract because she shares some of the demand risk by paying the deviation penalty for orders outside of the allowable range.

There are several ways in which the parties can adjust the terms of the percent deviation contract to satisfy the buyer's individual-rationality constraint. The supplier can offer the buyer a fixed transfer payment to share some of his gain. In some cases the supplier can offer a discounted wholesale price, w', that gives the buyer the same expected profit as she would attain under the traditional wholesale-price contract. We present a numerical examples using the discounted wholesale price strategy in Section 5. The remainder of this section illustrates the methodology required to find the requisite discounted wholesale price.

The A.∞ infinite capacity model is comparable to the traditional newsvendor, wholesale-price (NV) contract, since the supplier chooses and has the capacity to satisfy the entire order. The buyer's expected profit function under the NV contract is given by $\Pi^B_{NV} = (r-w)\int_0^\infty xf(x)dx$. Notice that this function is not dependent on any decision by the supplier, because under a traditional contract in this setting the buyer places orders for exactly the number of units needed with no demand risk. Comparing this expected profit to that under the percent deviation contract in (15) with the $q_1^{A.\infty}$ decision, the contracting parties wish to find w' such that $\Pi^B_{NV}(w) \geq \Pi^B_{A.\infty}(w')$, to ensure that the buyer earns at least as much expected profit under the percent deviation contract as she does under the original wholesale-price contract. We find that the discounted wholesale price given by

$$w' \leq \left(w - p\left[\frac{\int_0^{(1-d)q_1^{A.\infty}} ((1-d)q_1^{A.\infty} - x)f(x)dx + \int_{(1+d)q_1^{A.\infty}}^\infty (x-(1+d)q_1^{A.\infty})f(x)dx}{\int_0^\infty xf(x)dx} \right] \right)^+ \quad (18)$$

satisfies the buyer's participation constraint. The term in brackets represents the percentage of time that the deviation penalty will be paid. Consequently, the supplier must provide an allowance for this expected penalty if the buyer is to realize the same expected profit as in the newsvendor contract. If the right side of (18) assumes the value of zero, there is no discounted wholesale price mechanism that can satisfy the buyer's rationality constraint with the given contract parameters.

4.2 Channel Coordination

Supply chain research has shown that the total supply chain profit is maximized by a centralized firm making decisions that are best for the system as a whole. One main objective of supply chain contracts is to align each entity's own incentives to induce decentralized decisions that attain the maximal centralized supply chain profit. This achievement is commonly referred to as "channel coordination." We first examine the performance of the centralized channel and then develop mechanisms to coordinate the channel.

4.2.1 Centralized Channel Benchmark

In terms of a centralized channel, the buyer and the seller are viewed as a single entity trying to maximize its own expected profit. Hence, there is no wholesale price (w) paid from the sales department (the buyer) to the manufacturing department (the seller), and the penalties levied under the percent deviation contract (p and α) are not valid. The buyer's decisions are not relevant either since the single company does not order from itself; the combined firm must only determine the number of units to acquire early and the number to expedite.

If the cost structure for the centralized channel is such that $r - c_2 > -\beta$, the firm will satisfy additional demand beyond the number of pre-acquired goods up to capacity M. In this case the number of units to be expedited is given by $t_2 = (\min\{X - t_1, M\})^+$. The channel's expected profit function is

$$\begin{aligned}\Pi_{C.I.} &= r\left[\int_0^{t_1+M} xf(x)dx + (t_1+M)\left(1-F(t_1+M)\right)\right] + v\int_0^{t_1}(t_1-x)f(x)dx - c_1 t_1 - \\ &\quad c_2\left[\int_{t_1}^{t_1+M}(x-t_1)f(x)dx + M\left(1-F(t_1+M)\right)\right] - \beta\int_{t_1+M}^{\infty}(x-t_1-M)f(x)dx. \quad (19)\end{aligned}$$

This newsvendor-type profit function is concave, so first order optimality conditions show that the optimal solution for t_1 is $t_1^{C.I.} \in \left\{t : F(t+M) = \frac{r+\beta-c_1-(c_2-v)F(t)}{r+\beta-c_2}\right\}$.

If $r - c_2 < -\beta$, the loss from expediting or subcontracting to meet the marginal demand is larger than the cash outlay from the penalty paid to the customer for not satisfying its demand. Accordingly, the centralized channel will not expedite at the higher cost c_2; formally, we have $t_2 = 0$. The channel expected profit function now becomes

$$\Pi_{C.II.} = r\left[\int_0^{t_1} xf(x)dx + t_1(1-F(t_1))\right] + v\int_0^{t_1}(t_1-x)f(x)dx - c_1t_1 - \beta\int_{t_1}^{\infty}(x-t_1)f(x)dx. \quad (20)$$

The profit function in (20) is very similar to (19), except the expected revenue has been adjusted to reflect the fact that the centralized channel will not satisfy any demand more than t_1. The optimal number of units to acquire early is given by $t_1^{C.II.} \in \left\{t : F(t_1) = \frac{r+\beta-c_1}{r+\beta-v}\right\}$.

4.2.2 Infinite Expediting Capacity Channel Coordination

We first consider the infinite expediting capacity scenario where the shipper orders the entire demand, denoted by A.∞ above. The total supply chain profit, obtained by adding together the individual buyer and supplier profit functions, is equivalent to the centralized profit function for case C.I. given in (19) when $M = \infty$. This shows that any parameter values, as long as they are consistent with the inequalities defining these two cases, produce the maximum centralized profit. The percent deviation parameters (p and d), therefore, are merely a mechanism for profit distribution between the two parties in this case.

Coordinating the infinite capacity scenario in which the buyer does not order the full demand if it is greater than the upper limit of the deviation range is complex. This situation does not clearly correspond to either of the centralized scenarios, since they are differentiated by whether or not it is beneficial for the centralized firm to meet the excess demand and since the centralized decision maker never has such a penalty applied to partial orders. In this case the parties can achieve the centralized profit by setting the deviation penalty such that $p < r - w + \beta$. This causes the contract to shift to case A.∞ and, therefore, to the centralized case C.I. as well.

4.2.3 Finite Expediting Capacity Channel Coordination

Since the subgame-perfect Nash Equilibria for the scenarios in which the supplier has finite expediting capacity have a complicated form, different mechanisms are required for each possible decision pair. Consequently, we consider one possible decision pair show how the system can be coordinated given that particular decision. The procedure described below is applicable to all other possible decision pairs and case scenarios.

We consider Case B, in which the buyer orders the entire demand but the supplier chooses not to expedite, where the corresponding decision pair is $\left(q_1^{B.III.}, t_1^{B.III.}\right)$. The following lemma contains the channel coordinating condition for this case, which also applies in Case A when the same decision pair is optimal.

Lemma 2 *The decentralized channel in scenario B (and A) in which the SPNE decision pair is* $\left(q_1^{B.III.}, t_1^{B.III.}\right)$ *will be coordinated if the contract parameters are set such that*

$$\alpha + p + w = r + \beta. \tag{21}$$

The left-hand side of (21) comprises parameters that represent payments between the buyer and the supplier. These are set during contract negotiations as opposed to the right-hand side, which only contains parameters that we assumed were exogenous to the contract because they involve an outside party to the contract (the buyer's customer). The parties can coordinate the channel by setting α, p, and w according to (21).

4.3 Comparison to Quantity Flexibility Contracts

Since the percent deviation contract provides the buyer with order flexibility around an initial order estimate, it is constructive to compare its channel performance with the quantity flexibility contract, which affords the buyer similar flexibility. Tsay (1999) establishes that the quantity flexibility contract cannot coordinate the supply chain when the buyer is not bound by a minimum purchase commitment. The percent deviation contract, on the other hand, does coordinate the channel without establishing a floor on the buyer's order.

Let us consider analysis for a particular case, e.g., the B scenario in which the SPNE is $\left(q_1^{B.III.}, t_1^{B.III.}\right)$. Recall that this scenario can be coordinated by setting $\alpha+p+w = r+\beta$. To compare the quantity flexibility and percent deviation contracts, we need to analyze them in a similar framework. We apply the basic quantity flexibility contract structure but modify as follows to correspond to the percent deviation decision environment. We assume that the buyer's actual order in the quantity flexibility contract is made after the customer demand has been realized, as in the percent deviation scenario. Consequently, the supplier commits to fulfilling a maximum of t_1 units. The buyer establishes a minimum purchase commitment of $(1-\delta)q_1$ units when she provides the initial order estimate, q_1. If the buyer ends up ordering more units than she ultimately requires to satisfy the realized demand (as a result of the minimum purchase quantity), she receives u dollars per unit as a salvage value.

We assume that leftover units of inventory are no more valuable to the buyer than they are to the supplier (i.e., $u \leq v$). This is practical for several reasons. While it is true that goods generally appreciate in value as they move downstream in a supply chain, the buyer is not physically performing additional functions to add value to the product; consequently, the actual sale price of the salvaged product should be no higher than that which the supplier could receive if he sold it in the secondary market. Leftover product should be more valuable to the supplier in terms of expected revenue since he could likely use the product to fulfill demand from another buyer while the buyer may have limited outlets to offload the extra product. This is especially true in the market for truckload transportation, which was an inspiration for the percent deviation contract. Carriers would obviously place more value on an unassigned truck than any one particular shipper might.

Following the same backward-induction methodology we used in identifying the other equilibria, Theorem 4 provides the equilibrium decision for the quantity flexibility contract.

Theorem 4 *The SPNE decisions for the quantity flexibility contract are*

$$\left(q_1^{QF}, t_1^{QF}\right) = \begin{cases} \left(\frac{F^{-1}\left(\frac{r-w+\beta-\alpha}{r-v+\beta-\alpha}\right)}{1-\delta}, F^{-1}\left(\frac{r-w+\beta-\alpha}{r-v+\beta-\alpha}\right)\right), & \text{if} \quad (c_1-v)(r+\beta-\alpha)+wv-c_1u \geq \\ & (w+\alpha)(w-u) \\ \left(\{q|F((1-\delta)q)=0\}, F^{-1}\left(\frac{w+\alpha-c_1}{w+\alpha-v}\right)\right), & \text{otherwise.} \end{cases} \tag{22}$$

Suppose the parameter values are such that the quantity flexibility equilibrium decisions are the first pair in (22). We can write the expected total supply chain profit as the sum of the agents' individual expected profit functions, which reduces to

$$\begin{aligned} \Pi_{QF}^{SC} &= r\left[\int_0^{t_1^{QF}} xf(x)dx + t_1^{QF}\left(1-F\left(t_1^{QF}\right)\right)\right] + u\int_0^{t_1^{QF}}\left(t_1^{QF}-x\right)f(x)dx - c_1t_1^{QF} - \\ &\quad \beta\int_{t_1^{QF}}^{\infty}\left(x-t_1^{QF}\right)f(x)dx. \end{aligned} \tag{23}$$

Note that if $u = v$, for any value of t_1 we have $\Pi^{SC}_{QF}(t_1) = \Pi^{SC}_{C.II.}(t_1)$, where $\Pi^{SC}_{C.II.}(t_1)$ is the centralized supply chain profit in (20).

Theorem 5 *The percent deviation contract coordinates the supply chain in the following cases where the quantity flexibility contract fails to coordinate:*

i. When the salvage value is higher at the supplier ($u < v$), there are cases in which the centralized supply chain profit under the percent deviation contract always exceeds that attainable from the quantity flexibility contract.

ii. When the salvage values are equal for both parties ($u = v$), channel coordination efforts for quantity flexibility require either setting $\alpha < 0$ or $w < c_1$, both of which violate the underlying assumptions of the model.

In other supply chain contracting structures such as revenue-sharing agreements, it is possible for suppliers to benefit by selling goods for a wholesale price below their marginal cost of production as the second part of Theorem 5 requires. This strategy is successful because the supplier is receiving part of the buyer's revenue in addition to the wholesale price that he charges the buyer. Looking at the supplier's expected profit function under the quantity flexibility contract in (24), the supplier can either obtain w or v for each of the t_1 units he pre-acquires in advance of the buyer's order. If each of these values are less than c_1, he cannot earn positive expected profit by selling below his marginal cost.

We have thus shown that there are cases in which the quantity flexibility contract cannot coordinate the supply chain, while the percent deviation contract is able to achieve coordinated performance. The main difficulty the quantity flexibility contract has in this decision environment is that it establishes a minimum purchase commitment for the buyer. The percent deviation contract provides the buyer more flexibility by allowing them to choose to pay the penalties associated with ordering outside of the deviation range. Of course, in order to gain this flexibility, the contract must be more complex; therefore, the percent deviation contract would likely be more costly to manage in practice.

Table 2: Parameter declarations for numerical example

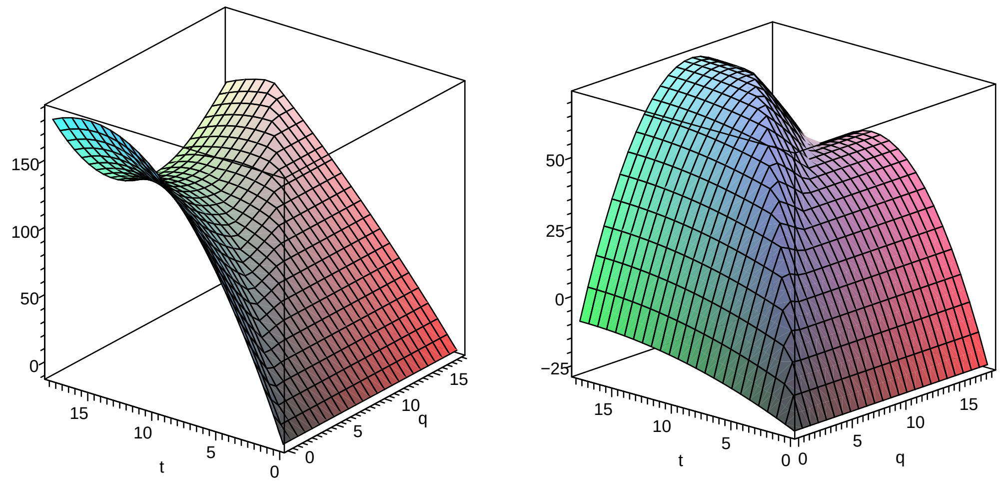
Parameter	Value	Parameter	Value	Parameter	Value
r	30	α	1	d	0.2
w	18	β	4	p	13
c_1	6	v	1	M	5
c_2	22				

5. Numerical Analysis

In this section we provide a numerical example that illustrates the behavior of the percent deviation contract as well as how parameters can be set to satisfy individual-rationality constraints and to coordinate the channel. We estimated the demand distribution used below from weekly shipping data provided by a major U.S. manufacturer. The demand random variable represents the number of shipments per week required from the supplier to a retailer on a particular origin-destination lane; based on the chi-squared goodness of fit test, the uniform distribution was appropriate for this data series. For the cost and contract parameters defined in Table 2, we constructed values that make relative sense in this manufacturer's business setting.

Since $r - w + \beta > p$ and $w + \alpha < c_2$, a percent deviation contract in this case would fall in scenario B., where the buyer orders the full demand and the supplier chooses not to expedite because it is too expensive. Under a traditional wholesale-price contract, the supplier pre-acquires 12.7059 units of inventory. The buyer and supplier expected profits are 95.54 and 76.24, respectively, resulting in a total supply chain expected profit of 171.78. If the firms acted as a centralized channel, the pre-acquisition amount would be 15.2727 with a total expected profit of 177.82.

Figures 2(a) and 2(b) depict the expected profit functions for the supplier and the buyer under a percent deviation contract in this example. We can apply the solution procedure for case B. to determine the SPNE decision pair of $(q_1^*, t_1^*(q_1^*)) = \left(q_1^{B.III.}, t_1^{B.III.}\right) = (10.3846, 15.0968)$, which results in expected profits of 71.53 and 106.26 for the buyer and the supplier, respectively, and a total supply chain expected profit of 177.79. Note that this decentralized percent deviation contract produces a supply chain profit very close to that



(a) (b)

Figure 2: Expected profit functions of (a) the supplier and (b) the buyer for the numerical example

of the centralized channel; this is due to the fact that the supplier's t_1 decision value is approximately equal to that of the centralized channel.

While the above percent deviation contract is close to coordinated as currently constructed, it does not satisfy the buyer's individual-rationality constraint when compared with the wholesale-price contract. Consequently, the percent deviation contract must be modified to give the buyer an incentive to accept it over the *status quo*. If the supplier offers a discounted wholesale price (as discussed in Section 4.1) of 15.2346, which represents an approximate discount of 15% off the original price of 18, the equilibrium decision pair becomes $(q_1^*, t_1^*(q_1^*)) = \left(q_1^{B.III.}, t_1^{B.III.}\right) = (10.3846, 14.1812)$. This contract results in expected profits of 95.54 and 82.08 for the buyer and the supplier, respectively, and a total supply chain expected profit of 177.62, which is still close to the centralized optimum of 177.82. This percent deviation contract with a discounted wholesale price satisfies the buyer's participation constraint and provides a higher profit for the supplier in relation to the traditional wholesale-price contract. Thus, we have shown how the percent deviation contract with a wholesale price discount can create a strictly Pareto-improving equilibrium over a traditional newsvendor mechanism.

6. Conclusions and Further Research

In this paper we have characterized the subgame-perfect Nash Equilibria of a dynamic supply chain game induced by the percent deviation contract, a mechanism that was motivated by our discussions with a major firm in the transportation industry. Due to the sequential extensive form of this supply chain game, many of the decisions are functions of those decisions made in earlier stages of the game. We also showed mechanisms for coordinating the decentralized channel, and we illustrated two methods to ensure that both parties' individual-rationality constraints are satisfied. The percent deviation contract can achieve channel coordination in some situations where the quantity flexibility contract fails to coordinate; this is mainly because the contract does not force the buyer to commit to a minimum purchase quantity. A numerical example based on a demand distribution estimated from industry data show that a properly-designed percent deviation contract can be strictly Pareto-improving.

The main result we have shown is that the percent deviation contract is a viable, albeit somewhat complicated, mechanism whereby the supplier can transfer some of his demand risk to the buyer. The prospect of receiving a deviation penalty for large or small buyer orders induces the supplier to pre-acquire more inventory than he ordinarily would, which increases the total capacity of the system. This extra ability to satisfy end-user demand benefits the entire system, enabling Pareto improvements for both parties.

Several trajectories exist for future research in this area. The first direction includes relaxing some of the assumptions that we made in these models. A natural extension would be adding some information asymmetry by including one party's proprietary information on costs or capacity. One could also extend the analysis by including nonlinear costs (reflecting production economies of scale) or some other pricing policy such as quantity discounts. More generally, future work incorporating dynamic decision environments could be useful, especially in multi-echelon supply chains. Comparison studies of various contracting mechanisms applied to the same scenario could lead to Pareto-improvements similar to the ones we found. Further analysis is also needed to incorporate the advanced demand information into operational production and transportation network models. Only then will the true value of the percent deviation contract be estimated for the system as a whole.

References

Balakrishnan, A., J. Geunes, M. F. Pangburn. 2004. Coordinating supply chains by controlling upstream variability propagation. *Manufacturing & Service Operations Management* **6**(2) 163–183.

Barnes-Schuster, D., Y. Bassok, R. Anupindi. 2002. Coordination and flexibility in supply contracts with options. *Manufacturing & Service Operations Management* **4**(3) 171–207.

Bazaraa, M. S., H. D. Sherali, C. M. Shetty. 1993. *Nonlinear Programming: Theory and Algorithms*. Second Edition. Wiley: New York.

Cachon, G. P. 2004. The allocation of inventory risk in a supply chain: Push, pull, and advance-purchase discount contracts. *Management Science* **50**(2) 222–238.

Cachon, G. P., M. Fisher. 2000. Supply chain inventory management and the value of shared information. *Management Science* **46**(8) 1032–1048.

Cachon, G. P., M. A. Lariviere. 2001. Contracting to assure supply: How to share demand forecasts in a supply chain. *Management Science* **47**(5) 629–646.

Donohue, K. L. 2000. Efficient supply contracts for fashion goods with forecast updating and two production modes. *Management Science* **46**(11) 1397–1411.

Drake, M. J. 2006. *The Design of Incentives for the Management of Supply and Demand.* Ph.D. Thesis. Georgia Institute of Technology.

Erkoc, M., S. D. Wu. 2005. Managing high-tech capacity expansion via capacity reservation. *Production & Operations Management* **14**(2) 232–251.

Ferguson, M. E., G. A. DeCroix, P. H. Zipkin. 2005. Commitment decisions with partial information updating. *Naval Research Logistics* **52**(8) 780–795.

Finley, F., S. Srikanth. 2005. Seven imperatives for successful collaboration. *Supply Chain Management Review* **9**(1) 30–37.

Gavirneni, S., R. Kapuscinski, S. Tayur. 1999. Value of information in capacitated supply chains. *Management Science* **45**(1) 16–24.

Huggins, E. L., T. L. Olsen. 2003. Supply chain management with guaranteed delivery.

Management Science **49**(9) 1154–1167.

Jin, M., S. D. Wu. 2001. Modeling capacity reservation in high-tech manufacturing. Working paper, Lehigh University.

Kulp, S. C., H. L. Lee, E. Ofek. 2004. Manufacturer benefits from information integration with retail customers. *Management Science* **50**(4) 431–444.

Lariviere, M. A., E. L. Porteus. 2001. Selling to the newsvendor: An analysis of price-only contracts. *Manufacturing & Service Operations Management* **3**(4) 293–305.

Lee, H. L. 2004. The triple-A supply chain. *Harvard Business Review* **82**(10) 102–112.

Lee, H. L., P. Padmanabhan, S. Whang. 1997. Information distortion in a supply chain: The bullwhip effect. *Management Science* **43**(4) 546–558.

Lee, H. L., K. C. So, C. S. Tang. 2000. The value of information sharing in a two-level supply chain. *Management Science* **46**(5) 626–643.

Parlar, M., Z. K. Weng. 1997. Designing a firm's coordinated manufacturing and supply decisions with short product life cycles. *Management Science* **43**(10) 1329–1344.

Taylor, T. A. 2002. Supply chain coordination under channel rebates with sales effort effects. *Management Science* **48**(8) 992–1007.

Tirole, J. 1988. *Theory of Industrial Organization.* MIT Press, Cambridge, MA.

Tsay, A. A. 1999. The quantity flexibility contract and supplier-customer incentives. *Management Science* **45**(10) 1339–1358.

Tsay, A. A., W. S. Lovejoy. 1999. Quantity flexibility contracts and supply chain performance. *Manufacturing & Service Operations Management* **1**(2) 89–111.

Tsay, A. A., S. Nahmias, N. Agrawal. 1999. Modeling supply chain contracts: A review. S. Tayur, R. Ganeshan, M. Magazine, eds. *Quantitative Methods for Supply Chain Management.* Kluwer Academic Publishers, Norwell, MA.

Weng, Z. K. 1997. Pricing and ordering strategies in manufacturing and distribution alliances. *IIE Transactions* **29**(8) 681–692.

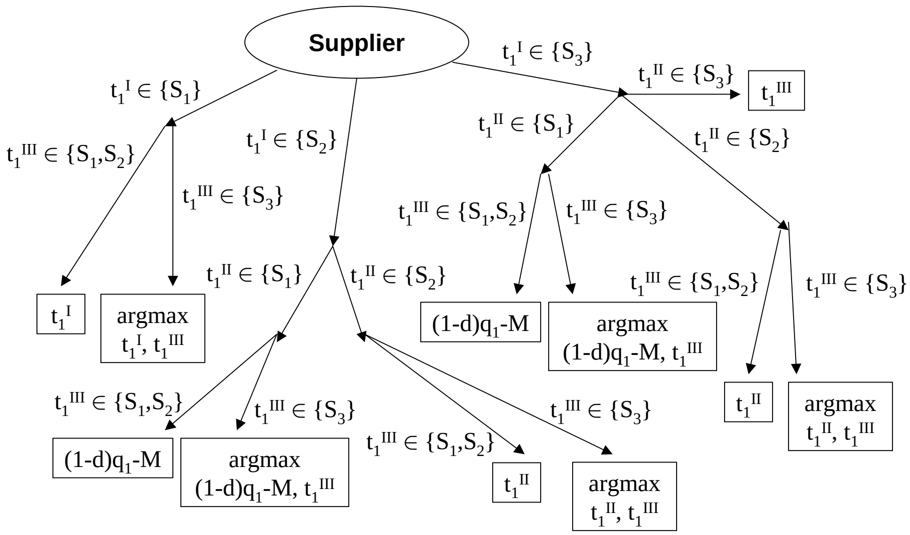


Figure 3: Supplier's best response decision tree for Case A

Appendix

Proof of Theorem 1. There are five possible values for the supplier's best response. Each of the three realizations of the supplier's expected profit function has an individual maximizer, shown in (3), (4), and (5). In addition, the two points where the pieces of the profit function converge ($t_1 = (1-d)q_1 - M$ and $t_1 = (1+d)q_1 - M$) are possible solutions. These solutions would occur when the maximizing t_1 values do not lie in their corresponding feasible regions. In order to establish the result in Theorem 1, we must first make some preliminary observations about the expected profit function that will help us later in the main proof.

Observation 2 *For all values of t_1 less than the lower boundary $((1-d)q_1-M)$, $\Pi^S_{A.II.}(t_1) > \Pi^S_{A.I.}(t_1)$ because the term representing the expected value of the lower deviation penalty paid is larger in $\Pi^S_{A.II.}$. For values of t_1 greater than the lower boundary, $\Pi^S_{A.II.}(t_1) < \Pi^S_{A.I.}(t_1)$.*

Observation 3 *For all values of t_1 less than the upper boundary $((1+d)q_1-M)$, $\Pi^S_{A.II.}(t_1) > \Pi^S_{A.III.}(t_1)$ because the term representing the expected value of the upper deviation penalty paid in $\Pi^S_{A.III.}$ is negative. For values of t_1 greater than the upper boundary, $\Pi^S_{A.II.}(t_1) < \Pi^S_{A.III.}(t_1)$.*

The supplier's best response function depends on the values of the three maximizers relative to the feasible boundaries. There are 27 possible cases because each of the three decisions can potentially lie in three regions; however, the following results show that several of these cases are not possible.

Lemma 3 *It is not possible to have* $t_1^{A.I.} < (1-d)q_1 - M < t_1^{A.II.}$.

Lemma 4 *It is not possible to have* $t_1^{A.III.} < (1+d)q_1 - M < t_1^{A.II.}$.

After using Lemmas 3 and 4 to reduce the number of possible cases, we can determine the overall best response for each given the values of the individual maximizers by applying Observations 1–3. Summarizing all of the scenarios, we obtain the solution given in (6) and depicted in Figure 3. ■

Proof of Proposition 1. We use the difference function to determine when the supplier chooses each t_1 for given values of q_1 in a range where his best response is known to be the maximizer of his expected profit from a set of two values. We then use this information to characterize explicitly the ranges of q_1 that induce each value of t_1. If the difference function is positive for a value of q_1, then the supplier will choose $t_1^{A.III.}$; he will select the other possible decision if the function is negative. The number of ranges for q_1 is determined by the number of changes of sign in the difference function. The difference function related to the decision in (8) is given by

$$\begin{aligned}
&\Delta\left(t_1^{A.I.}\right) = \\
&w\left[\int_{t_1^{A.I.}+M}^{t_1^{A.III.}+M} xf(x)dx + \left(t_1^{A.III.}+M\right)\left(1-F\left(t_1^{A.III.}+M\right)\right) - \left(t_1^{A.I.}+M\right)\left(1-F\left(t_1^{A.I.}+M\right)\right)\right] + \\
&p\left[\int_0^{(1-d)q_1}((1-d)q_1 - x)f(x)dx - \int_0^{t_1^{A.I.}+M}\left(t_1^{A.I.}+M-x\right)f(x)dx\right] + \\
&p\left[\int_{(1+d)q_1}^{t_1^{A.III.}+M}(x-(1+d)q_1)f(x)dx + \left(t_1^{A.III.}+M-(1+d)q_1\right)\left(1-F\left(t_1^{A.III.}+M\right)\right)\right] + \\
&+v\left[\int_0^{t_1^{A.III.}}\left(t_1^{A.III.}-x\right)f(x)dx - \int_0^{t_1^{A.I.}}\left(t_1^{A.I.}-x\right)f(x)dx\right] - c_1\left(t_1^{A.III.}-t_1^{A.I.}\right) - \\
&\alpha\left[\int_{t_1^{A.III.}+M}^{\infty}\left(x-t_1^{A.III.}-M\right)f(x)dx - \int_{t_1^{A.I.}+M}^{\infty}\left(x-t_1^{A.I.}-M\right)f(x)dx\right] - \\
&c_2\left[\int_{t_1^{A.III.}}^{t_1^{A.III.}+M}\left(x-t_1^{A.III.}\right)f(x)dx - \int_{t_1^{A.I.}}^{t_1^{A.I.}+M}\left(x-t_1^{A.I.}\right)f(x)dx\right] + \\
&c_2\left[M\left(F\left(t_1^{A.I.}+M\right) - F\left(t_1^{A.III.}+M\right)\right)\right].
\end{aligned}$$

This difference function is convex in q_1 since $\frac{\partial^2\Delta}{\partial q_1^2} = p(1-d)^2 f((1-d)q_1) + p(1+d)^2 f((1+d)q_1) > 0$. We can begin by evaluating the difference function at the two endpoints of the region defined in (8); that is, $q_1 = \frac{t_1^{A.I.}+M}{1-d}$ and $q_1 = \frac{t_1^{A.III.}+M}{1+d}$. If the difference function is positive for one value and negative for the other, then convexity implies that there exists a single threshold value of q_1 in the interval where the difference function changes sign. The buyer can use these supplier decision values to evaluate her best selection of q_1 in this region with respect to her expected profit function.

If the difference function is positive for both endpoint values of q_1, then it is possible that there are zero, one, or two points where the function switches sign. If there are zero or one switching points, then the supplier will choose $t_1 = t_1^{A.III.}$ for all values of q_1 in the region. If there are two switching points, then for values of q_1 between these two values, the supplier will choose $t_1 = t_1^{A.I.}$. He will choose $t_1 = t_1^{A.III.}$ for all other values of q_1. If the difference function is negative for both endpoint values, then convexity implies that it will be negative for all values of q_1; thus, the supplier will always choose $t_1 = t_1^{A.I.}$.

The difference function related to the decision in (10) is given by

$$\begin{aligned}
&\Delta\left(t_1^{A.II.}\right) = \\
&w\left[\int_{t_1^{A.II.}+M}^{t_1^{A.III.}+M} xf(x)dx + \left(t_1^{A.III.}+M\right)\left(1-F\left(t_1^{A.III.}+M\right)\right) - \left(t_1^{A.II.}+M\right)\left(1-F\left(t_1^{A.II.}+M\right)\right)\right] + \\
&p\left[\int_{(1+d)q_1}^{t_1^{A.III.}+M}(x-(1+d)q_1)f(x)dx + \left(t_1^{A.III.}+M-(1+d)q_1\right)\left(1-F\left(t_1^{A.III.}+M\right)\right)\right] + \\
&+v\left[\int_0^{t_1^{A.III.}}\left(t_1^{A.III.}-x\right)f(x)dx - \int_0^{t_1^{A.II.}}\left(t_1^{A.II.}-x\right)f(x)dx\right] - c_1\left(t_1^{A.III.}-t_1^{A.II.}\right) - \\
&\alpha\left[\int_{t_1^{A.III.}+M}^{\infty}\left(x-t_1^{A.III.}-M\right)f(x)dx - \int_{t_1^{A.II.}+M}^{\infty}\left(x-t_1^{A.II.}-M\right)f(x)dx\right] - \\
&c_2\left[\int_{t_1^{A.III.}}^{t_1^{A.III.}+M}\left(x-t_1^{A.III.}\right)f(x)dx - \int_{t_1^{A.II.}}^{t_1^{A.II.}+M}\left(x-t_1^{A.II.}\right)f(x)dx\right] + \\
&c_2\left[M\left(F\left(t_1^{A.II.}+M\right)-F\left(t_1^{A.III.}+M\right)\right)\right].
\end{aligned}$$

This difference function is convex in q_1 since $\frac{\partial^2\Delta}{\partial q_1^2} = p(1+d)^2 f((1+d)q_1) > 0$, and it is also decreasing in q_1 because $\frac{\partial\Delta}{\partial q_1} = -p(1+d)(1-F((1+d)q_1)) < 0$. This means that if the difference function is negative when $q_1 = \frac{t_1^{A.II.}+M}{1+d}$, which is the lower limit of the range defined in (10), then the supplier will always choose $t_1 = t_1^{A.II.}$. Likewise, if the difference function is positive at the upper endpoint of the range $\left(q_1 = \min\left\{\frac{t_1^{A.II.}+M}{1-d}, \frac{t_1^{A.III.}+M}{1+d}\right\}\right)$, then the supplier will always select $t_1 = t_1^{A.III.}$. If the difference function is positive for the lower endpoint and negative for the upper endpoint, then there exists exactly one point where the difference function changes sign, and we have two distinct ranges of q_1 values where the two t_1 decisions are chosen.

The difference function related to the decision in (9) is given by

$$\Delta\left((1-d)q_1-M\right)=$$
$$w\left[\int_{(1-d)q_1}^{t_1^{A.III.}+M} xf(x)dx+\left(t_1^{A.III.}+M\right)\left(1-F\left(t_1^{A.III.}+M\right)\right)-(1-d)q_1\left(1-F\left((1-d)q_1\right)\right)\right]+$$
$$p\left[\int_{(1+d)q_1}^{t_1^{A.III.}+M}(x-(1+d)q_1)f(x)dx+\left(t_1^{A.III.}+M-(1+d)q_1\right)\left(1-F\left(t_1^{A.III.}+M\right)\right)\right]+$$
$$+v\left[\int_0^{t_1^{A.III.}}\left(t_1^{A.III.}-x\right)f(x)dx-\int_0^{(1-d)q_1-M}\left((1-d)q_1-M-x\right)f(x)dx\right]-$$
$$c_1\left(t_1^{A.III.}-(1-d)q_1+M\right)-\alpha\left[\int_{t_1^{A.III.}+M}^{\infty}\left(x-t_1^{A.III.}-M\right)f(x)dx-\int_{(1-d)q_1}^{\infty}\left(x-(1-d)q_1\right)f(x)dx\right]-$$
$$c_2\left[\int_{t_1^{A.III.}}^{t_1^{A.III.}+M}\left(x-t_1^{A.III.}\right)f(x)dx-\int_{(1-d)q_1-M}^{(1-d)q_1}\left(x-(1-d)q_1+M\right)f(x)dx\right]+$$
$$c_2\left[M\left(F\left((1-d)q_1\right)-F\left(t_1^{A.III.}+M\right)\right)\right].$$

Here one of the potential supplier decisions is an explicit function of the buyer's q_1 decision, so the difference function is more complex. Specifically, the function is not necessarily convex or concave. For a given set of parameters, then the exact switching points can be determined by simple numerical search methods. In many realizations the difference function will be well-behaved; thus, a similar analysis to that performed for the previous two cases above would suffice for these situations. ■

Proof of Theorem 4. We will solve for the subgame-perfect Nash equilibrium decisions under a quantity flexibility contract via backward induction. The parameters in the B. scenario are such that the buyer orders $q_2^*=\max\{X,(1-\delta)q_1\}$, where X denotes the realized customer demand. The supplier's expected profit can thus be written as

$$\begin{aligned}\Pi_{QF}^S &= w\left[\int_0^{(1-\delta)q_1}(1-\delta)q_1 f(x)dx+\int_{(1-\delta)q_1}^{t_1} xf(x)dx+t_1(1-F(t_1))\right]+\\ & v\left[\int_0^{(1-\delta)q_1}(t_1-(1-\delta)q_1)f(x)dx+\int_{(1-\delta)q_1}^{t_1}(t_1-x)f(x)dx\right]-c_1t_1-\\ & \alpha\int_{t_1}^{\infty}(x-t_1)f(x)dx. \end{aligned} \qquad (24)$$

Since the supplier's expected profit function is concave, first-order optimality conditions imply that the supplier's optimal decision is $t_1=F^{-1}\left(\frac{w+\alpha-c_1}{w+\alpha-v}\right)$. There is one additional consideration, though, since the buyer is guaranteed to order at least $(1-\delta)q_1$. The supplier should pre-acquire at least the minimum purchase quantity because he is guaranteed to sell it. Thus, the supplier's optimal decision is $t_1^{QF}=\max\left\{(1-\delta)q_1,F^{-1}\left(\frac{w+\alpha-c_1}{w+\alpha-v}\right)\right\}$.

The buyer's expected profit function is given by

$$\Pi_{QF}^B = r\left[\int_0^{t_1^{QF}} xf(x)dx+t_1^{QF}(1-F(t_1^{QF}))\right]-$$

$$w\left[\int_0^{(1-\delta)q_1}(1-\delta)q_1 f(x)dx+\int_{(1-\delta)q_1}^{t_1^{QF}} xf(x)dx+t_1^{QF}(1-F(t_1^{QF}))\right]+$$
$$u\int_0^{(1-\delta)q_1}((1-\delta)q_1-x)f(x)dx+(\alpha-\beta)\int_{t_1^{QF}}^{\infty}(x-t_1^{QF})f(x)dx. \tag{25}$$

We can solve for the buyer's optimal decision, as before, by assuming that the supplier's decision takes on each of the two possible values and then optimizing the buyer's profit subject to the constraint that makes the supplier's decision valid. The decision pair $\left(q_1^{QF},t_1^{QF}\right)=\left(\frac{F^{-1}\left(\frac{r-w+\beta-\alpha}{r-v+\beta-\alpha}\right)}{1-\delta},F^{-1}\left(\frac{r-w+\beta-\alpha}{r-v+\beta-\alpha}\right)\right)$ is optimal if $F^{-1}\left(\frac{r-w+\beta-\alpha}{r-v+\beta-\alpha}\right)\geq F^{-1}\left(\frac{w+\alpha-c_1}{w+\alpha-v}\right)$, which reduces to $(c_1-v)(r+\beta-\alpha)+wv-c_1u\geq(w+\alpha)(w-u)$. If this inequality is reversed, $\left(q_1^{QF},t_1^{QF}\right)=\left(\frac{F^{-1}\left(\frac{w+\alpha-c_1}{w+\alpha-v}\right)}{1-\delta},F^{-1}\left(\frac{w+\alpha-c_1}{w+\alpha-v}\right)\right)$. In this case the supplier's decision is fixed regardless of the value of q_1^{QF}, so the buyer can reduce her demand risk by offering $q_1^{QF}\in\{q|F((1-\delta)q)=0\}$ such that there is no probability of customer demand below the minimum purchase amount. ■

Proof of Theorem 5. If $u<v$, then clearly $\Pi_{QF}^{SC}\leq\Pi_{C.II.}^{SC}$ for every value of t_1, and there exist some values of t_1 where the inequality is strict. Consequently, coordination is not possible in these cases because leftover goods are less valuable in the buyer's possession, which is where they reside under quantity flexibility.

Now let $u=v$. If $t_1^{QF}=t_1^{C.II}$, then $\Pi_{QF}^{SC}=\Pi_{C.II.}^{SC}$, and we would have a coordinated supply chain. Thus, we want to have $\frac{r-w+\beta-\alpha}{r-v+\beta-\alpha}=\frac{r+\beta-c_1}{r+\beta-v}$. Since $\alpha(\geq 0)$, the penalty the supplier pays the buyer for not satisfying units she orders, is the one parameter over which the contracting parties are assumed to have control under quantity flexibility, we solve for the coordinating condition

$$\alpha=\frac{(r+\beta-v)(c_1-w)}{c_1-v}. \tag{26}$$

Examining the components of (26) individually, we see that the first term in the numerator is greater than zero because $r>v$ and $\beta\geq 0$, as is the denominator. So if $w>c_1$ by our initial assumption, then this would require a negative α. We could have a positive coordinating α if we allowed the supplier to sell the goods below cost. ■

Supplemental Appendix for "Facilitating Demand Risk-Sharing with the Percent Deviation Contract"

Proof of Lemma 1. We will define the three realizations of (2) as follows:

$$\begin{aligned}
\Pi^S_{A.I.} &= w\left[\int_0^{t_1+M} xf(x)dx + (t_1+M)\left(1-F(t_1+M)\right)\right] + p\int_0^{t_1+M}\left(t_1+M-x\right)f(x)dx + \\
& v\int_0^{t_1}(t_1-x)f(x)dx - c_1t_1 - c_2\left[\int_{t_1}^{t_1+M}(x-t_1)f(x)dx + M\left(1-F(t_1+M)\right)\right] - \\
& \alpha\int_{t_1+M}^{\infty}(x-t_1-M)f(x)dx \qquad (27)\\
\Pi^S_{A.II.} &= w\left[\int_0^{t_1+M} xf(x)dx + (t_1+M)\left(1-F(t_1+M)\right)\right] + p\int_0^{(1-d)q_1}\left((1-d)q_1-x\right)f(x)dx + \\
& v\int_0^{t_1}(t_1-x)f(x)dx - c_1t_1 - c_2\left[\int_{t_1}^{t_1+M}(x-t_1)f(x)dx + M\left(1-F(t_1+M)\right)\right] - \\
& \alpha\int_{t_1+M}^{\infty}(x-t_1-M)f(x)dx \qquad (28)\\
\Pi^S_{A.III.} &= w\left[\int_0^{t_1+M} xf(x)dx + (t_1+M)\left(1-F(t_1+M)\right)\right] + p\int_0^{(1-d)q_1}\left((1-d)q_1-x\right)f(x)dx + \\
& p\left[\int_{(1+d)q_1}^{t_1+M}(x-(1+d)q_1)f(x)dx + \left(t_1+M-(1+d)q_1\right)\left(1-F(t_1+M)\right)\right] + \\
& v\int_0^{t_1}(t_1-x)f(x)dx - c_1t_1 - c_2\left[\int_{t_1}^{t_1+M}(x-t_1)f(x)dx + M\left(1-F(t_1+M)\right)\right] - \\
& \alpha\int_{t_1+M}^{\infty}(x-t_1-M)f(x)dx. \qquad (29)
\end{aligned}$$

The second derivative of (27) taken with respect to t_1 is $(p+c_2-w-\alpha)f(t_1+M)-(c_2-v)f(t_1)$, which is negative for all values of t_1 if $w+\alpha > p+c_2$ based on the parameter conditions of scenario A. The second derivative of (28) is $(c_2-w-\alpha)f(t_1+M)-(c_2-v)f(t_1)$, and the second derivative of (29) is $(c_2-w-\alpha-p)f(t_1+M)-(c_2-v)f(t_1)$. Both of these expressions are negative for all values of t_1 without the extra condition. ■

Proof of Lemma 2. This case can easily be compared with the centralized case C.II. in which the centralized firm also does not expedite. The total expected supply chain profit for the voluntary compliance case is

$$\Pi^{SC}_{B.} = r\left[\int_0^{t_1} xf(x)dx + t_1(1-F(t_1))\right] + v\int_0^{t_1}(t_1-x)f(x)dx - c_1t_1 - \beta\int_{t_1}^{\infty}(x-t_1)f(x)dx. \qquad (30)$$

Comparing (30) with the centralized supply chain profit in (20), it is easily seen that the two profits will be equal if the t_1 decisions are equal, which is accomplished if $\frac{w+\alpha-c_1+p}{w+\alpha-v+p} = \frac{r+\beta-c_1}{r+\beta-v}$. Simplifying this equality yields the channel coordinating condition. ■

Proof of Lemma 3. This result follows directly from Lemma 5. ■

Proof of Lemma 4. Assume that this relationship is true. Since the piecewise functions are concave from Lemma 1, $t_1^{A.III.}$ is the single maximum of $\Pi^S_{A.III.}$, and $\Pi^S_{A.III.}(t_1)$ is decreasing for values of $t_1 > t_1^{A.III.}$. Consequently, $\Pi^S_{A.III.}(t_1^{A.III.}) > \Pi^S_{A.III.}((1+d)q_1-M) > \Pi^S_{A.III.}(t_1^{A.II.})$. Since $\Pi^S_{A.II.}((1+d)q_1-M) = \Pi^S_{A.III.}((1+d)q_1-M)$ from Observation 1, we have $\Pi^S_{A.III.}(t_1^{A.III.}) > \Pi^S_{A.II.}((1+d)q_1-M) > \Pi^S_{A.III.}(t_1^{A.II.})$. Observation 3 states that $\Pi^S_{A.III.}(t_1^{A.II.}) > \Pi^S_{A.II.}(t_1^{A.II.})$, which implies $\Pi^S_{A.III.}(t_1^{A.III.}) > \Pi^S_{A.II.}((1+d)q_1-M) > \Pi^S_{A.III.}(t_1^{A.II.}) > \Pi^S_{A.II.}(t_1^{A.II.})$. The statement $\Pi^S_{A.II.}((1+d)q_1-M) > \Pi^S_{A.II.}(t_1^{A.II.})$ contradicts the result that $t_1^{A.II.}$ maximizes $\Pi^S_{A.II.}$. ■

Lemma 5 *If $w+\alpha > p+c_2$, then $t_1^{A.I.} \geq t_1^{A.II.}$.*

Proof. Suppose, on the contrary, $t_1^{A.I.} < t_1^{A.II.}$, which implies that $F(t_1^{A.I.}+M) \leq F(t_1^{A.II.}+M)$. Substituting the values given in (3) and (4), we have

$$\begin{aligned}
\frac{w+\alpha-c_1-(c_2-v)F(t_1^{A.I.})}{w+\alpha-c_2-p} &\leq \frac{w+\alpha-c_1-(c_2-v)F(t_1^{A.II.})}{w+\alpha-c_2} \\
(c_2-v)(w+\alpha-c_2)\left[F(t_1^{A.II.})-F(t_1^{A.I.})\right] &\leq -p\left(w+\alpha-c_1-(c_2-v)F(t_1^{A.II.})\right). \quad (31)
\end{aligned}$$

The left side of (31) is positive, and the right side is negative since the numerator in (4) must be positive at $t_1^{A.II.}$. (The denominator is positive due to the parameter relationship defining case A.) This leads to a contradiction. ■

Lemma 6 *If $w+\alpha > p+v$, then $t_1^{I.B.II.} \leq \min\left\{t_1^{B.I.}, t_1^{B.III.}\right\}$.*

Proof. We follow the same contradiction procedure as in the proof of Lemma 5. First, assume $t_1^{B.II.} > t_1^{B.I.}$, which implies that $F\left(t_1^{B.II.}\right) \geq F\left(t_1^{B.I.}\right)$. Further substitution yields

$$\begin{aligned}
\frac{w+\alpha-c_1}{w+\alpha-v} &\geq \frac{w+\alpha-c_1}{w+\alpha-v-p} \\
-p(w+\alpha-c_1) &\geq 0,
\end{aligned}$$

which leads to a contradiction since $w+\alpha-c_1 \geq 0$.

Similarly, we assume that $t_1^{B.II.} > t_1^{B.III.}$; consequently, $F\left(t_1^{B.II.}\right) \geq F\left(t_1^{B.III.}\right)$. We obtain the following contradiction through substitution.

$$\frac{w+\alpha-c_1}{w+\alpha-v} \geq \frac{w+\alpha-c_1+p}{w+\alpha-v+p} \Rightarrow v \geq c_1,$$

which is a contradiction since $c_1 > v$. ■

Proof of Lemma 7. We will define the four functions resulting from (7) as follows:

$$\begin{aligned}
\Pi_{A.I.}^{B} &= (r-w)\left[\int_0^{t_1^*(q_1)+M} xf(x)dx + \left(t_1^*(q_1)+M\right)\left(1-F\left(t_1^*(q_1)\right)\right)\right] - \\
&\quad p\int_0^{t_1^*(q_1)} \left(t_1^*(q_1)+M-x\right)f(x)dx + (\alpha-\beta)\int_{t_1^*(q_1)}^{\infty} \left(x-t_1^*(q_1)-M\right)f(x)dx \qquad (32)\\
\Pi_{A.II.}^{B} &= (r-w)\left[\int_0^{t_1^*(q_1)+M} xf(x)dx + \left(t_1^*(q_1)+M\right)\left(1-F\left(t_1^*(q_1)\right)\right)\right] - \\
&\quad p\int_0^{(1-d)q_1} \left((1-d)q_1-x\right)f(x)dx + (\alpha-\beta)\int_{t_1^*(q_1)}^{\infty} \left(x-t_1^*(q_1)-M\right)f(x)dx \qquad (33)\\
\Pi_{A.III.}^{B} &= (r-w)\left[\int_0^{t_1^*(q_1)+M} xf(x)dx + \left(t_1^*(q_1)+M\right)\left(1-F\left(t_1^*(q_1)\right)\right)\right] - \\
&\quad p\int_0^{(1-d)q_1} \left((1-d)q_1-x\right)f(x)dx + (\alpha-\beta)\int_{t_1^*(q_1)}^{\infty} \left(x-t_1^*(q_1)-M\right)f(x)dx - \qquad (34)\\
&\quad p\left[\int_{(1+d)q_1}^{t_1^*(q_1)+M} (x-(1+d)q_1)f(x)dx + \left(t_1^*(q_1)+M-(1+d)q_1\right)\left(1-F\left(t_1^*(q_1)+M\right)\right)\right]. \\
\Pi_{A.IV.}^{B} &= (r-w)\left[\int_0^{(1-d)q_1} xf(x)dx + (1-d)q_1\left(1-F\left((1-d)q_1\right)\right)\right] - \\
&\quad p\int_0^{(1-d)q_1} \left((1-d)q_1-x\right)f(x)dx + (\alpha-\beta)\int_{(1-d)q_1}^{\infty} \left(x-(1-d)q_1\right)f(x)dx \qquad (35)
\end{aligned}$$

The second derivative of (32) is zero, so this component function is concave. The second derivative of (33) equals $-p(1-d)^2 f((1-d)q_1)$, which is negative for all values of q_1; thus, this function is concave. Similarly, the second derivative of (34) is $-p(1-d)^2 f((1-d)q_1) - p(1+d)^2 f((1+d)q_1)$, which is negative for all q_1. The second derivative of (35) is $-(r-w-\alpha+\beta)(1-d)^2 f((1-d)q_1) - p(1-d)^2 f((1-d)q_1)$, which is also negative for all q_1, making this function concave as well. ■

Proof of Theorem 2. The following lemma establishes the piecewise-concavity of (7) with respect to the four realizations of the buyer's expected profit function.

Lemma 7 *The buyer's expected profit function realizations resulting from (7) are concave.*

Since the buyer's four individual profit function realizations are concave from Lemma 7, we can use the KKT conditions to solve for the optimal q_1 for each function over the region of q_1 values where the function is valid, as defined in 1.

We first maximize (32) over the region $q_1 \geq \frac{t_1^{A.I.}+M}{1-d}$. Since (32) is not dependent on q_1, any value $q_1^{A.I.} \equiv \left\{q : q \geq \frac{t_1^{I.A.I.}+M}{1-d}\right\}$ is optimal.

We consider (33) over the region $q_1 \leq \frac{t_1^{A.II.}+M}{1-d} \bigcap q_1 \geq \frac{t_1^{A.II.}+M}{1+d} \bigcap q_1 \geq \frac{t_1^{A.III.}+M}{1+d}$. Taking the partial derivative and setting it equal to zero yields $p(1-d)F((1-d)q_1) = 0$. Since only the lower deviation penalty exists in this profit function realization, the buyer wants to make her initial order estimate as small as possible to avoid paying the penalty. Consequently, $q_1^* = q_1^{A.II.} \equiv \max\left\{\frac{F^{-1}(0)}{1-d}, \frac{\max\left\{t_1^{A.II.}, t_1^{A.III.}\right\}+M}{1+d}\right\}$.

We want to maximize (34) over the region $q_1 \leq \frac{t_1^{A.II.}+M}{1+d} \bigcap q_1 \leq \frac{t_1^{A.III.}+M}{1+d}$. First order optimality conditions yield $q_1^* = q_1^{A.III.} \equiv \{q : (1+d) = (1-d)F((1-d)q)+(1+d)F((1+d)q)\}$, which is feasible if it is smaller than $\frac{\min\left\{t_1^{A.II.}, t_1^{A.III.}\right\}+M}{1+d}$.

Finally, we maximize (35) over the region $q_1 \geq \frac{t_1^{A.II.}+M}{1-d} \bigcap q_1 \leq \frac{t_1^{A.I.}+M}{1-d} \bigcap q_1 \geq \frac{t_1^{A.III.}+M}{1+d}$. The first order conditions give us $q_1^* = q_1^{A.IV.} \equiv \left\{q : F((1-d)q_1) = \frac{r-w-\alpha+\beta}{r-w-\alpha+\beta+p}\right\}$, which is feasible if $\max\left\{\frac{t_1^{A.II.}+M}{1-d}, \frac{t_1^{A.III.}+M}{1+d}\right\} \leq q_1^{\beta} \leq \frac{t_1^{A.I.}+M}{1-d}$. ■

Proof of Theorem 3. The proof of this result follows the same logic as that of Theorem 1, utilizing the results from Lemma 6. ■