A Note on Air-Cargo Capacity Contracts

Kannapha Amaruchkul∗ • William L. Cooper† • Diwakar Gupta†

∗School of Applied Statistics
National Institute of Development Administration
118 SeriThai Road
Bangkapi, Bangkok 10240, Thailand

†Program in Industrial and Systems Engineering
University of Minnesota
111 Church Street S.E.
Minneapolis, MN 55455, USA

December 11, 2009
A Note on Air-Cargo Capacity Contracts

Abstract

Carriers (airlines) use medium-term contracts to allot bulk cargo capacity to forwarders who deliver consolidated loads for each flight in the contractual period (season). Carriers also sell capacity to direct-ship customers on each flight. We study capacity contracts between a carrier and a forwarder when certain parameters such as the forwarder’s demand, operating cost to the carrier, margin, and reservation profit are its private information. We propose contracts in which the forwarder pays a lump sum in exchange for a guaranteed capacity allotment and receives a refund for each unit of unused capacity according to a pre-announced refund rate. We obtain an upper bound on the informational rent paid by the carrier for a menu of arbitrary allotments and identify conditions under which it can eliminate the informational rent and induce the forwarder to choose the overall optimal capacity allotment (i.e. one that maximizes the combined profits of the carrier and the forwarder).

Keywords: Air Cargo Revenue Management, Capacity Contracts, Asymmetric Information, Informational Rents.
1 Introduction

Carriers (airlines) and freight forwarders are partners in the air-cargo service chain. Carriers own airplanes and operate flights on which they may transport cargo. Forwarders purchase bulk cargo capacity by contracting with airlines, but typically do not themselves own planes. Forwarders sell the purchased capacity to individual shippers, consolidate shipments into larger units, and deliver consolidated loads to the carriers. For international cargo, forwarders also provide certain value-adding services that carriers are not well equipped to perform on their own, such as customs clearance and export-import documentation. More than 60% of domestic and 90% of international air-cargo tonnage is tendered by forwarders (Coppersmith 2003), underscoring the importance of capacity contracts in the air-cargo industry.

Carriers sell cargo space via two different mechanisms: preallocation and ad hoc sales (see Bazaraa et al. 2001 and Gupta 2008 for details). Prior to the start of each season, carriers allocate cargo capacity to forwarders on that season’s flights. (A season is a pre-designated length of time, ranging from a few weeks to a year.) Capacity sold to a forwarder in this manner is called an allotment (Kasilingam 1996). Carriers also sell space to customers on an ad hoc or space-available basis without prior commitments. In our terminology, ad hoc demand is comprised of demand from forwarders who are not regular customers of the carrier, demand from forwarders who are regular customers but who have loads that exceed their allotment, and spot sales to those who are not forwarders. In this paper, we focus on the interactions between a single carrier and a single forwarder.

How should the carrier choose the combination of allotments and financial terms offered to the forwarder? To answer this question, we utilize the Principal-Agent framework where we cast the carrier as the principal and the forwarder as the agent. In such a framework, the first-best profit is the maximum combined profit of the principal (carrier) and agent (forwarder) if a single decision maker were to coordinate their decisions. We say that the carrier can achieve the first-best profit when the sum of the carrier’s and the forwarder’s profits equals the first-best profit and the forwarder earns its reservation profit. Reservation profit is the minimum amount that the forwarder needs to earn in order to enter into a contract with the carrier.

Salanié (1997), Section 2.2.3, shows that when there is no hidden information and the
principal has unlimited capacity, then a contract consisting of two elements — a capacity allocation and a lump-sum payment — is sufficient for the principal to achieve the first-best profit. In our setting, this is true even though the carrier’s capacity is limited and it also serves demand from ad hoc customers. There are other pairs of parameters that can achieve the first-best profit for the carrier, e.g. a freight rate for capacity sold and a refund rate for returns of unused capacity. Such contracts are frequently proposed in the supply chain literature (see, e.g. Pasternack 1985 and Emmons and Gilbert 1998).

In the presence of informational asymmetry, the problem is significantly more difficult. Laffont and Martimort (2002), p. 36-44, and Salanié (1997), p. 21-26, assume specific forms of the agent’s payoff and show that when offering a menu of two-parameter contracts, each consisting of a capacity allocation and a lump-sum payment, the principal cannot achieve the first-best profit in the presence of hidden information. The presence of private information typically allows the party with more information (the forwarder in our setting) to collect informational rent, that is, to earn more than its reservation profit (which it obtains under full information, as indicated above). One may view the carrier as paying the rent to the forwarder, because the rent earned by the forwarder comes at the expense of the carrier.

In this note, we develop a formulation for the study of contracts with three parameters: (1) an allotment, (2) a lump-sum payment required by the carrier to guarantee the allotment to the forwarder, and (3) a refund rate that determines the amount that the forwarder receives for the unused portion of the allotment. Our aim is to shine light on the magnitude of informational rent that the carrier would have to pay and to identify a set of sufficient conditions under which the carrier can achieve the first-best profit in the presence of informational asymmetry. There are several reasons for considering contracts with refunds. First, offering refunds is a common practice — carriers are known to allow forwarders to return unwanted space a few days before a flight departs and receive a refund (Hellermann 2006). Second, refunds provide an incentive to the forwarder to return unused space in a timely fashion, which is then added to the pool of available ad hoc space thereby reducing the carrier’s uncertainty about the amount of space available for ad hoc sales. Finally, as we show in this note, the provision of refunds also allows the carrier to earn the first-best profit under certain conditions.
In our formulation, the carrier offers a menu consisting of the aforementioned three-parameter contracts to the forwarder. The forwarder then either chooses one of the offered contracts or rejects the carrier’s proposition. The forwarder’s private information is captured by a type, which determines the forwarder’s customer demand, margin, and reservation profit. The forwarder’s private information may also include induced operating cost to the carrier because typically the carrier does not know exactly what kind of cargo the forwarder will book. The carrier bases its contracting decision on a prior probability distribution over a finite type space, reflecting its beliefs about the forwarder’s type. When certain forwarder characteristics (demand distribution, margin, reservation profit) are monotone in the type index, we obtain an upper bound on the informational rent that the carrier must pay. We also show that the carrier can achieve the first-best profit when, in addition, the forwarder’s margin and reservation profit are common knowledge. Our formulation can also be used to obtain the carrier’s best course of action when the first-best profit is not attainable, although this may require numerical computations.

Forwarders can typically offer more services than carriers, and also offer lower prices to shippers in comparison to carriers’ standard tariffs because they receive a volume discount from the carriers (Sigworth 2004). Consequently, shippers prefer to use forwarders unless they have an emergency shipment, or shipments containing perishable or hazardous materials (Thuemer 2005). Similarly, ad hoc demand from forwarders consists of shipments that exceed their allotments. Typically such shipments become known to the carriers very close to the flight’s departure. For these reasons, we assume that ad hoc shipments arrive after the forwarder returns the unused portion of its allotment, if any.

In certain segments of the worldwide air cargo markets, carriers have little pricing power. It has been noted in the literature that they are sometimes unable to collect an up-front payment and that the forwarders pay only for the capacity they use (Bazaraa et al. 2001). Clearly, in such instances, carriers would be unable to achieve the first-best profit and may end up paying substantial informational rents. Depending on the constraints imposed on the principal (i.e. on the principal’s ability to choose contract parameters and their values), a variety of less-profitable solutions are possible. Our model benefits a carrier by identifying a possible contract structure that it should strive for in its negotiations with the forwarder.
Because the three-parameter contracts studied in this note lead to an optimal capacity allocation, there may exist opportunities for sharing the associated excess profits in a manner that benefits both actors. Such a division could be achieved through a bargaining process.

Next, we present a brief review of related literature. Riordan and Sappington (1988) identify a sufficient condition under which the first-best profit is attainable under asymmetric information. In their setting, there exists an exogenous random signal that is correlated with the agent’s type, and contracts (struck prior to realization of the signal) specify the financial flow between the principal and agent as a function of the realized value of the signal. Hence, different signal realizations yield different financial flows. In our setting, the amount of returned capacity may be viewed as a signal because it is correlated with the forwarder’s type and affects financial flows (via the refund for unused capacity). However, our model is different from Riordan and Sappington’s because the value of the signal also depends on the menu of offered allotments. In particular, the forwarder decides how much capacity to return based on the allotment, the refund rate, and the amount it can earn from shippers.

Crémer and McLean (1985) provide a condition similar to Riordan and Sappington’s in a multiple-agent setting, where the agents’ types are correlated. Makowski and Mezzetti (1994) assume that the agents’ types are independent, but that the principal, who pays a lump sum to each agent, must make sure that the sum of these payments is zero. Karabuk and Wu (2005) apply the approach in Makowski and Mezzetti (1994) to achieve the first-best profit in a semiconductor capacity allocation problem.

Capacity management models that are not specific to air-cargo operations (see Van Mieghem 2003 for a review) have several features in common with the air-cargo literature, but there are also significant differences. For example, in both cases, the capacity owner guarantees allotments to some buyers in return for reservation fees. Capacity is also sold to other customers without guaranteeing availability in advance. A key difference between the two types of models is that in the air-cargo setting, the owner (i.e. the carrier) does not know what type of cargo will be booked by the buyer (i.e. the forwarder) and the operational costs induced by the buyer are therefore private. In contrast, capacity management literature concerns contracts for a particular product or service for which the owner knows its operational costs and typically private information is either the demand or the cost parameters of the
agent. Another difference is that in the air-cargo setting, it is a common practice for the buyer to return the unused portion of its allotment before the departure of each flight.

Examples of related work in other settings (not carrier-forwarder interactions) can be found in Ha (2001), Corbett et al. (2004), Lovejoy (2006), and Burnetas et al. (2007). In these articles, a contract has two parameters, which are the equivalents of the allotment and the up-front payment in our formulation. They compare how common two-parameter contracts perform under informational asymmetry. As mentioned earlier, two-parameter contracts are not sufficient to realize the first-best profit.

Özer and Wei (2006) study capacity commitment decisions of a supplier that relies on demand forecast information from a manufacturer. Much of their analysis deals with two-parameter contracts. However, they do briefly discuss three-parameter contracts. It can be deduced from their Theorem 8 that contracts that specify a wholesale price, a lump-sum payment, and a capacity reservation level allow the supplier in their setting to achieve the first-best profit when the only private information is the manufacturer’s demand. Although the models developed in their paper and ours are quite different — for example, Özer and Wei assume that the manufacturer’s demand realization is a linear function of its type but we require no particular functional relationship — their result is analogous to our Corollary 1, which states that under some conditions, the carrier can achieve the first-best when the forwarder’s private information consists of its demand.

Prior work on carrier-forwarder contracts (e.g. Hellermann 2006 and Gupta 2008) assumes symmetric information. Thus, a key distinction of our work is that we allow the forwarder to possess private information. Gupta (2008) studies the carrier-forwarder incomplete contract problem, where the forwarder’s non-verifiable effort level determines the magnitude of its demand. We do not consider incomplete contracts within the framework of asymmetric information. Such models are possible topics for future research.

The main contribution of this note is the development and analysis of a model for evaluating carrier-forwarder capacity contracts when the forwarder has private information. Specifically, we derive an upper bound on the informational rent that the carrier would pay on account of the forwarder’s private information for a given menu of capacity allotments. We also identify a set of allotment-independent sufficient conditions that allow the carrier to
achieve the first-best profit.

The remainder of this paper is organized as follows. Section 2 contains a principal-agent model formulation for our problem. In Section 3, we provide bounds on the informational rent and identify a set of conditions under which the first-best profit is attainable. Finally, Section 4 contains concluding remarks. Proofs are in the Appendix.

2 Formulation

The forwarder, one of $n$ possible types in a type space $\mathcal{N} = \{1, \ldots, n\}$, knows its type, but the carrier views the forwarder’s type as a random variable $T$ with prior distribution $\{\theta_i = P(T = i) : i \in \mathcal{N}\}$, where $\theta_i > 0$ for all $i \in \mathcal{N}$. The forwarder’s type determines its demand distribution, margin, reservation profit, and the variable operating cost incurred by the carrier from handling the forwarder’s shipments. Let $p_i > 0$ (resp, $p_0 > 0$) be a type-$i$ forwarder’s (carrier’s) margin per unit of sales (ad hoc sales), and $c_i \in [0, p_i]$ be the carrier’s operating cost per unit of allotment used by a type-$i$ forwarder.

The carrier has $\kappa$ units of cargo capacity per flight, which we assume to be a one-dimensional fixed quantity. In practice, individual cargo booking decisions may take into consideration packages’ weights, volumes, and shapes. In addition, capacity may be random, and depend upon weather conditions, the number of passengers, their baggage, and the amount of fuel on the plane. The focus of this article is not on making operational booking decisions (see Popescu et al. 2006, Amaruchkul et al. 2007, and Levin et al. 2008 for papers dealing with operational issues), but rather on medium-term capacity contracts. At the time of negotiating such contracts, neither the forwarder nor the carrier knows the exact nature of shipments they will receive in the future. Moreover, factors that make capacity random are also difficult to forecast. Therefore, freight rates and contract negotiations are based on weight with a dimensional adjustment and the carriers make contract decisions based on expected capacity. In summary, it is appropriate to focus, as we do in this article, on one dimension (weight) and not to model the randomness of capacity.

Our model considers a single flight. In practice, capacity contracts can last several months (referred to as a season) and are comprised of multiple repeat flights. It is possible to modify our analysis to include multiple flights in the season. Our results carry over — at the
expense of more complicated notation — to settings where demands for space on the flights are conditionally identically distributed (given the forwarder’s type) and spillover of demand from one flight to another is not significant. We expect the latter to hold because air-cargo shipments are time sensitive. In the event of an oversold situation, carriers/forwarders are more likely to find space on another carrier’s flight or send the shipment via an alternate route on one of the carrier’s own flights. This would be preferable to waiting for the next repeat flight of the carrier, which is typically scheduled for another day.

The carrier offers a menu of potential contracts to the forwarder. Each contract specifies an allotment $x \in \mathcal{X} := [0, \kappa]$ of cargo capacity that the carrier grants the forwarder, a fixed non-negative payment $f \geq 0$ from the forwarder to the carrier for the allotment, and a strictly positive refund rate $r > 0$ to the forwarder for each unit of the allotment it returns unused to the carrier. The forwarder then chooses one of the contracts or else rejects all offers. If the latter occurs, the forwarder can still earn a positive profit from serving its customers by different means. The net profit that can be earned in this way is the forwarder’s reservation profit. An alternate interpretation of reservation profit is that it is the forwarder’s fixed cost of doing business with the carrier. We do not model the forwarder’s recourse in detail, instead we assume that the reservation profit of a type-$i$ forwarder is known to be $\epsilon_i \geq 0$.

If the forwarder chooses one of the contracts, then that contract is implemented during the season. Once the contract is in place and the forwarder has paid $f$ to the carrier, events occur as follows: (1) the forwarder’s demand realizes and the forwarder makes bookings, (2) the forwarder returns any unused portion of the allotment to the carrier at a pre-negotiated time (usually 24 to 48 hours) prior to the flight’s departure, (3) the carrier’s ad hoc demand realizes and the carrier makes bookings, and (4) the flight departs with the booked cargo. We do not allow the refund rate to be zero (or negative) because in such situations, the forwarder would have no incentive to return unused capacity prior to the flight’s departure.

Let $D$ and $D_0$ be non-negative random variables that represent, respectively, demand to the forwarder and direct-ship (ad hoc) demand to the carrier. The joint distribution of $(D, D_0)$ depends upon the type of the forwarder. For $i \in \mathcal{N}$, let $G_i(\cdot)$ be the conditional distribution function of $D$, given that the forwarder is type-$i$; that is, $G_i(y) = P(D \leq y | T = i)$ for $y \in \mathbb{R}_+$. Similarly, we use $H_i(\cdot)$ to denote the conditional distribution function of $D_0$;
i.e., \( H_i(y) = P(D_0 \leq y | T = i) \) for \( y \in \mathbb{R}_+ \). To simplify notation, let \((D_i, D_{i0})\) denote random variables whose joint distribution satisfies \( P(D_i \leq y, D_{i0} \leq y') = P(D \leq y, D_0 \leq y' | T = i) \); hence, \( P(D_i \leq y) = P(D \leq y | T = i) = G_i(y) \) and \( P(D_{i0} \leq y') = P(D_0 \leq y' | T = i) = H_i(y') \). To simplify the exposition we further assume that \( G_i : \mathbb{R}_+ \rightarrow [0,1] \) and \( H_i : \mathbb{R}_+ \rightarrow [0,1] \) have densities \( g_i(\cdot) \) and \( h_i(\cdot) \) and that each \( G_i(\cdot) \) and \( H_i(\cdot) \) is strictly increasing on \([0,\kappa]\). (Strict monotonicity is not assumed over \((\kappa,\infty)\).) We use a “bar” atop a distribution function to denote its complement; e.g. \( \bar{H}_i(y) = 1 - H_i(y) \).

Let \( C' = \{(f,r) : f \geq 0, r > 0\} \) be the set of possible (fixed payment, refund rate)-pairs, and let \( C = \{(x,f,r) : x \in \mathcal{X}, (f,r) \in C'\} \) be the set of possible contracts. Let \( \pi_i(x,f,r) \) denote the conditional expected (net) profit of a forwarder that picks contract \((x,f,r) \in C\), given that it is type \( i \). To develop an expression for \( \pi_i(x,f,r) \) we consider cases with \( r \geq p_i \) and \( r < p_i \) separately. It is important to consider the possibility that \( r \geq p_i \) because in the presence of information asymmetry, it may benefit the carrier to allow some forwarder types to earn a positive profit without selling any capacity so that it can earn more profit from other forwarder types. Below, the term “booking limit” refers to a quantity such that the amount of cargo booked is the smaller of demand and the booking limit.

If \( r \geq p_i \) then it is in the best interest of the forwarder to book nothing (which can be viewed as the forwarder setting a booking limit of zero) and to return the entire allotment to obtain a refund of \( rx \). (When \( r = p_i \), the forwarder’s profit does not depend upon its booking policy, because the forwarder can return each unsold unit of its allotment for a refund that is equal to its per-unit margin. In such situations, we assume that the forwarder books nothing, as this requires the least “effort” on the part of the forwarder.) On the other hand, if \( r < p_i \), then it is in the best interest of the forwarder to book as much as possible (which corresponds to the forwarder setting a booking limit of \( x \)), in which case it will earn \( p_i \min(x, D_i) \) from bookings and return the unused portion of its allotment to earn an additional \( r(x - D_i)^+ \). Note that when \( r < p_i \) the total net payments under a contract \((x,f,r)\) are the same as those that would occur if (i) there is no up-front payment and the forwarder receives allotment \( x \), (ii) the forwarder agrees to pay the carrier \( f/x \) per unit of the \( \min(x, D_i) \) units of cargo it sends to the carrier, and (iii) the forwarder agrees to pay the carrier \( f/x - r \) per unit for the portion \((x - D_i)^+\) of the allotment \( x \) it does not use.
Although our analysis does not consider overbooking, it is worth pointing out that if we were to allow the forwarder to overbook, then we would see similar behavior. That is, if \( r \geq p_i \), then the forwarder would book nothing. In contrast, if \( r < p_i \), then the forwarder would set a booking limit — say \( b \geq x \) — resulting in \( \min(b, D_i) \) bookings.

Taking conditional expectations and accounting for the fixed payment, we have

\[
\pi_i(x, f, r) = \begin{cases} 
  rx - f & \text{if } r \geq p_i \\
  v_i(x) + rE(x - D_i)^+ - f & \text{if } r < p_i,
\end{cases}
\]

where

\[
v_i(x) = p_i E[\min(x, D_i)] = p_i \int_0^\infty \min(x, y)g_i(y)dy = p_i \int_0^x \bar{G}_i(y)dy
\]

is the forwarder’s conditional expected net revenue (sometimes called contribution) from sales, given that the forwarder is type \( i \in \mathcal{N} \). When \( r_i \in (0, p_i) \), we can simplify (1) as follows.

\[
\pi_i(x, f, r) = v_i(x) + rz_i(x) - f = p_i x - (p_i - r)z_i(x) - f,
\]

where

\[
z_i(x) = E(x - D_i)^+ = x - \int_0^x \bar{G}_i(u)du
\]

is the expected amount of capacity returned by a type-\( i \) forwarder that receives allotment \( x \). This representation of \( \pi_i(x, f, r) \) is useful for proving results presented in the next section.

Similarly, we define \( w_i(\ell) \) to be the carrier’s conditional expected contribution from direct-ship demand, net of handling costs from the forwarder’s bookings, when the forwarder sets a booking limit of \( \ell \);

\[
w_i(\ell) = E[p_0 \min\{D_{i0}, \kappa - \min(\ell, D_i)\} - c_i \min(\ell, D_i)].
\]

In (5) above, the term \( \min\{D_{i0}, \kappa - \min(\ell, D_i)\} \) represents the carrier’s direct-ship bookings. Hence, expression (5) assumes, as explained in the introduction, that the carrier finalizes capacity sales to direct shippers only after knowing how much of the allotment is used by the forwarder. Note that \( w_i(0) = p_0 E[\min(D_{i0}, \kappa)] \geq w_i(x) \) for any \( x \in \mathcal{X} \).

Given that the forwarder is type \( i \), the conditional expected profit of the carrier that
awards contract \((x, f, r) \in \mathcal{C}\) is

\[
\rho_i(x, f, r) = \begin{cases} 
  w_i(0) - rx + f & \text{if } r \geq p_i \\
  w_i(x) - rE(x - D_i)^+ + f & \text{if } r < p_i.
\end{cases}
\]  

(6)

The expressions in (6) reflect the facts that if \(r \geq p_i\) then the forwarder will set its booking limit to zero, and if \(r < p_i\) then forwarder will set its booking limit equal to its allotment. For \(x > 0\), the carrier’s profit function \(\rho_i(\cdot)\) defined in (6) has a discontinuity at \(r = p_i\), at which point the forwarder switches from booking as much as possible to booking nothing. This discontinuity complicates our analysis in Section 3.

3 Results

To set the stage for our main analysis, we shall first consider settings in which the carrier knows the forwarder’s type, and chooses a contract to maximize its own conditional expected profit. It will be convenient for our subsequent developments to write the optimization problem in two stages, with an inner optimization over fixed payments and refund rates, and an outer optimization over allotments. For a type-\(i\) forwarder, the carrier’s benchmark problem without hidden information is the following:

\[
\text{Problem } \mathcal{B}_i \quad \zeta_i = \sup_{x_i \in \mathcal{X}} \zeta_i(x_i) \tag{7}
\]

where

\[
\zeta_i(x_i) = \sup_{(f_i, r_i) \in \mathcal{C}'} \rho_i(x_i, f_i, r_i) \tag{8}
\]

s.t. \(\pi_i(x_i, f_i, r_i) \geq \epsilon_i\).  

(9)

Constraint (9) ensures that a type-\(i\) forwarder earns at least its reservation profit. Such constraints are called individual rationality (IR) constraints.
Let

$$\psi_i(x) = v_i(x) + w_i(x)$$  \hspace{1cm} (10)$$

be the combined expected profit of the carrier and the forwarder when the forwarder books up to $x$, conditional upon the forwarder being type $i$. To understand this interpretation of $\psi_i(x)$ defined in (10), observe that if $r \in (0, p_i)$, then $\pi_i(x, f, r) + \rho_i(x, f, r) = v_i(x) + w_i(x) = \psi_i(x)$. If $r \geq p_i$, then $v_i(0) = 0$ and $\pi_i(x, f, r) + \rho_i(x, f, r) = w_i(0) = \psi_i(0)$. The function $\psi_i(x)$ is identical to the objective function of the two-class passenger revenue management problem studied in Brumelle et al. (1990) if we were to view the forwarder’s demand as demand for the discount fare of $p_i - c_i$, direct-ship demand as demand for the full fare $p_0$, and $x$ as the booking limit on discount tickets. We let $x^*_i$ denote a maximizer of $\psi_i(x)$; i.e., $x^*_i \in \arg\max_{x \in X} \psi_i(x)$.

If the carrier decides not to do business with the forwarder, then by using all its capacity to book direct-ship cargo, the carrier can earn $p_0E[\min(D_{i0}, \kappa)] = w_i(0) = \psi_i(0)$. Hence, the carrier will prefer to do business with the forwarder if and only if $\psi_i(0) < \zeta_i$. Taking this into account, the profit of the carrier is $\max\{\zeta_i, \psi_i(0)\}$.

Given a type $i \in \mathcal{N}$, the first-best profit is the combined profit of the carrier and the forwarder that would result if there were a single decision maker that coordinated the carrier’s and forwarder’s actions to maximize their combined profit. If the decision maker grants an allotment of $x_i$ to the forwarder from the carrier, then the combined profit will be $\psi_i(x_i)$. If the decision maker does not grant an allotment to the forwarder, then the carrier earns $\psi_i(0)$ and the forwarder earns $\epsilon_i$, for a combined total of $\psi_i(0) + \epsilon_i$. Hence, the first-best profit is $\max\{\psi_i(x^*_i), \psi_i(0) + \epsilon_i\}$. We say that the first best is attainable if the carrier earns $\max\{\psi_i(x^*_i) - \epsilon_i, \psi_i(0)\}$ and the forwarder earns $\epsilon_i$; that is, if carrier and forwarder together earn the first-best profit and the forwarder earns its reservation profit. We say that the first best is attainable in Problem $\mathbb{B}_i$ if the first best is attainable and $\zeta_i > \psi_i(0)$. The qualifier “in Problem $\mathbb{B}_i$” reflects the fact that under the stated condition, the carrier and forwarder enter into a contract specified by Problem $\mathbb{B}_i$.

Let $x^*_i$ be such that $p_i x^*_i = \epsilon_i$. The following lemma gives the solution of Problem $\mathbb{B}_i$ when $D_0$ and $D$ are conditionally independent given $T = i$. As shown in the proof, if $\epsilon_i > 0$ and the forwarder receives an allotment less than $x^*_i$ and a refund rate less than its margin, then
the forwarder’s profit from bookings and returned capacity will be less than its reservation profit. For such an allotment, the carrier would need to offer the forwarder a refund rate that meets or exceeds the forwarder’s margin to satisfy the IR constraint in Problem $B_i$.

**Lemma 1.** Suppose that $D_0$ and $D$ are conditionally independent given $T = i$. Then, the first best is attainable and the unique maximizer of $\psi_i(x)$ over $x \in \mathcal{X}$ is

$$x^*_i = \left[ \kappa - \bar{H}^{-1}_i \left( \frac{p_i - c_i}{p_0} \right) \right]^+$$

(11)

where we define $\bar{H}^{-1}_i((p_i - c_i)/p_0)) = 0$ if $p_i - c_i > p_0$. In Problem $B_i$ we have

$$\zeta_i = \begin{cases} 
\psi_i(x^*_i) - \epsilon_i & \text{if } p_i x^*_i > \epsilon_i \\
\max\{\psi_i(0), \psi_i(x^*_i)\} - \epsilon_i & \text{if } p_i x^*_i \leq \epsilon_i < p_i \kappa \\
\psi_i(0) - \epsilon_i & \text{if } p_i \kappa \leq \epsilon_i.
\end{cases}$$

(12)

The first best is attainable in Problem $B_i$ if the inequality $\psi_i(0) < \psi_i(x^*_i) - \epsilon_i$ also holds.

It follows from the preceding lemma that restricting the lump-sum payment $f_i$ to be non-negative in Problem $B_i$ causes no disadvantage to the carrier, because the carrier cannot do better than to attain the first best.

Note that the forwarder’s demand distribution does not appear in (11), which can be derived by differentiating $\psi_i(x)$ to obtain the first-order optimality condition, $(p_i-c_i)\bar{G}_i(x) - p_0\bar{H}_i(\kappa-x)\bar{G}_i(x) = 0$ for $x \in [0, \kappa]$. If $D_0$ and $D$ are not conditionally independent given $T = i$, then (11) generally will not maximize $\psi_i(x)$. In that case, a maximizer can be found by a one-dimensional search on $[0, \kappa]$. We can relax the assumption of conditional independence in the lemma and still attain the first best; the key observation is that the carrier can offer a contract with an allotment that maximizes $\psi_i(x_i)$ and choose the lump-sum payment to make the IR constraint tight.

Observe that when $r_i < p_i$, the total amount the forwarder pays to the carrier for its allotment must be no more than the most the forwarder can earn $(p_i x^*_i)$ from the allotment; otherwise, the forwarder would not do business with the carrier and make its non-negative reservation profit. Note also that if instead the carrier sold the amount $x^*_i$ to a type-$i$ forwarder as ad hoc bookings it would charge $(p_0 + c_i)x^*_i$, which is more than $p_i x^*_i$ if $p_0 + c_i > p_i$. 

12
Hence, in this sense, the formulation captures volume discounts offered by the carrier to the forwarder.

We next make a few comments on two-parameter contracts. As mentioned earlier, we require that the refund rate be strictly positive so that it is in the interest of the forwarder to return the unused portion of its allotment in advance of the flight’s departure. If we ignore this issue for the time being and suppose (1) and (6) are valid with \( r = 0 \), then we can study two-parameter contracts by requiring that \( r_i = 0 \) in Problem \( \mathbb{B}_i \). With such an added requirement, arguments similar to those above show that the first best is attainable with two-parameter contracts as well. This result is consistent with previous work on two parameter contracts without informational asymmetry mentioned in the Introduction section.

We now turn to the setting in which the carrier does not know the forwarder’s type. The carrier’s problem is to design a menu of contracts to maximize its expected profit. By the Revelation Principle (see, e.g., Laffont and Martimort 2002 or Salanié 1997), this problem can be formulated assuming that the forwarder truthfully reports its type to the carrier, who then assigns a contract from the menu to the forwarder based upon the report. In the resulting optimization problem (Problem A below), the triplet \( (x_i, f_i, r_i) \) is a decision variable that represents the contract awarded to a forwarder that reports its type to be \( i \). Let \( x = (x_1, \ldots, x_n) \).

\[
\text{Problem A} \quad \xi = \sup_{x \in \mathbb{X}^n} \xi(x) \tag{13}
\]

where

\[
\xi(x) = \sup_{\{(f_i, r_i) \in \mathbb{C} : i \in \mathbb{N}\}} \sum_{i=1}^{n} \theta_i \rho_i(x_i, f_i, r_i) \tag{14}
\]

\[
\text{s.t. } \pi_i(x_i, f_i, r_i) \geq \pi_i(x_j, f_j, r_j) \quad \text{for all } i, j \in \mathbb{N} \tag{15}
\]

\[
\pi_i(x_i, f_i, r_i) \geq \epsilon_i \quad \text{for all } i \in \mathbb{N}. \tag{16}
\]

The incentive compatibility (IC) constraints (15) ensure that the forwarder reports its true type. The IR constraints (16) guarantee that the forwarder’s conditional expected profit is at least its reservation profit \( \epsilon_i \), given that it is type \( i \). It will be useful to introduce
non-negative slack variables \( \{ \delta_i \geq 0 : i \in \mathcal{N} \} \) and write the IR constraints (16) as equalities:

\[
\pi_i(x_i, f_i, r_i) = \epsilon_i + \delta_i \quad \text{for all } i \in \mathcal{N}.
\]

The slack variable \( \delta_i \) represents the amount by which the profit of a type-\( i \) forwarder exceeds its reservation profit. Recall from Lemma 1 that under symmetric information, the forwarder earns its reservation profit \( \epsilon_i \) when demands are conditionally independent. Hence, if \( \xi > \sum_{i=1}^{n} \theta_i \psi_i(0) \) (so that the two parties enter into a contract) and demands are conditionally independent, then \( \delta_i \) represents the amount of informational rent that a type-\( i \) forwarder can extract from the carrier as a result of informational asymmetry.

Note that we again require that the lump-sum payments be non-negative. In general, allowing negative lump-sum payments may improve the optimal value \( \xi \); however, it seems unlikely that a carrier would ever offer such a contract in practice as it would amount to paying the forwarder to take capacity. Nevertheless, the results below remain unchanged if we allow the lump-sum payments to be negative.

In Problem A we have implicitly assumed that if a type-\( i \) forwarder is indifferent between reporting its type to be either \( i \) or \( j \neq i \) [i.e. if \( \pi_i(x_i, f_i, r_i) = \pi_i(x_j, f_j, r_j) \) so that the \((i, j)\)-th IC constraint is binding], then it will report its type to be \( i \). Similarly, if a type-\( i \) forwarder can earn only \( \epsilon_i \), i.e. if the \( i \)-th IR constraint is binding, then it will do business with the carrier. Many papers that use the Revelation Principle to solve mechanism design problems rely on such assumptions; see, e.g., Laffont and Martimort (2002, p. 37). As noted by Laffont and Martimort, these issues can be resolved by imposing strict inequalities in the IC and IR constraints, which results in an “arbitrarily small” loss of profit to the principal.

Under Assumptions A1 and A2 that follow, Theorem 1 below provides an lower bound on the supremum in (14) for a fixed vector of allotments \( \hat{x} = (\hat{x}_1, \cdots, \hat{x}_n) \). Without loss of generality we hereafter index the forwarder types such that \( \hat{x}_1 \leq \hat{x}_2 \leq \cdots \leq \hat{x}_n \). The theorem does not require an assumption of conditional independence of \( D_0 \) and \( D \).

**A1** Certain forwarder characteristics are monotone in type. In particular, (i) \( G_1(y) \leq G_2(y) \leq \cdots \leq G_n(y) \) for all \( y \in \mathbb{R} \); (ii) \( 0 < p_1 \leq p_2 \leq \cdots \leq p_n \); and (iii) \( \epsilon_1 \geq \epsilon_2 \geq \cdots \geq \epsilon_n \geq 0 \).
(i) \( p_i \hat{x}_1 > \epsilon_1 \), and (ii) for all \( i, j \in \mathcal{N} \) such that \( i < j \) we have

\[
(p_j - p_i) \int_0^{\hat{x}_i} \bar{G}_j(u) \, du \leq \sum_{k=i}^{j-1} (p_{k+1} - p_k) \int_0^{\hat{x}_k} \bar{G}_{k+1}(u) \, du \leq (p_j - p_i) \int_0^{\hat{x}_j} \bar{G}_i(u) \, du. \quad (17)
\]

Assumption A1(i) states that the forwarder’s demand is stochastically increasing in its type; see, e.g., Shaked and Shanthikumar (1994) for a review of stochastic orders. Assumption A2(i) states that if a type-1 forwarder is able to book its entire allotment, then its gross profit (without the fixed payment) exceeds its reservation profit. Assumptions A1(ii), A1(iii), and A2(i) together imply that \( p_i \hat{x}_i > \epsilon_i \) for all \( i \in \mathcal{N} \). Assumption A2(ii) is used below to enforce IC constraints.

Theorem 1. Consider allotments \( \hat{x} = (\hat{x}_1, \ldots, \hat{x}_n) \) such that Assumptions A1 and A2 hold. Then \( \xi(\hat{x}) \geq \sum_{i=1}^n \theta_i [\psi_i(\hat{x}_i) - \epsilon_i - \bar{\delta}_i] \) where \( \bar{\delta}_1 = 0 \) and

\[
\bar{\delta}_i = \sum_{k=1}^{i-1} \left[ \int_0^{\hat{x}_k} \left\{ (p_{k+1} - p_k) \bar{G}_{k+1}(u) - (p_k - p_1) \bar{G}_k(u) \right\} \, du \right] + \epsilon_1 - \epsilon_i \quad \text{for } i = 2, \ldots, n.
\]

For each \( i \in \mathcal{N} \), the quantity \( \bar{\delta}_i \) in the theorem provides an upper bound on the informational rent, i.e., the amount that the carrier will fail to recover from a type-\( i \) forwarder due to asymmetric information.

In the corollary below, we apply the theorem when the forwarder’s margin and reservation profit are known (\( p_i \equiv p \) and \( \epsilon_i \equiv \epsilon \)). We will also use the following assumption, which is used to ensure that (11) gives the optimal allotments and that those allotments are increasing in the forwarder’s type.

A3 (i) \( D_0 \) and \( D \) are conditionally independent given \( T \); (ii) \( x_1^* \leq x_2^* \leq \cdots \leq x_n^* \) where \( x_i^* \) is defined by (11) for each \( i \in \mathcal{N} \).

If \( p_i \equiv p \) and also (a) \( \bar{H}_1(y) \geq \bar{H}_2(y) \geq \cdots \geq \bar{H}_n(y) \) for all \( y \in \mathbb{R} \) and (b) \( c_1 \geq c_2 \geq \cdots \geq c_n \), then Assumption A3(ii) holds. A simple sufficient condition that implies Assumption A3(i) as well as (a) is that the carrier’s demand \( D_0 \) is independent of the vector \( (D, T) \) of the forwarder’s demand and type, in which case \( H_i(y) \) does not depend upon \( i \), i.e., \( H_i(y) \equiv H(y) \).
Corollary 1. Suppose that Assumptions A1(i) and A3 hold, and that $p_1 = \cdots = p_n = p$ and $\epsilon_1 = \cdots = \epsilon_n = \epsilon$. Suppose also that $px^*_1 > \epsilon$. Then, $\xi = \sum_{i=1}^{n} \theta_i \zeta_i$. If, in addition, $\psi_i(x^*_i) - \epsilon > \psi_i(0)$ for all $i \in \mathcal{N}$, then $\xi = \sum_{i=1}^{n} \theta_i \max\{\psi_i(x^*_i) - \epsilon, \psi_i(0)\} > \sum_{i=1}^{n} \theta_i \psi_i(0)$.

We may interpret Corollary 1 as follows. Under the stated conditions, the first-best profit is attainable in Problem $A$ under asymmetric information if the first-best profit is attainable in Problem $B_i$ for each type $i \in \mathcal{N}$ of forwarder under symmetric information. If $\psi_i(x^*_i) - \epsilon > \psi_i(0)$ for all $i \in \mathcal{N}$, then the carrier does business with the forwarder because $\xi > \sum_{i=1}^{n} \theta_i \psi_i(0)$. Moreover, given that the forwarder is type $i$, the carrier earns the corresponding first-best profit $\psi_i(x^*_i) = \max\{\psi_i(x^*_i), \psi_i(0) + \epsilon\}$ less the forwarder’s reservation profit $\epsilon_i$, and the forwarder is left with its reservation profit $\epsilon_i$. This occurs even though the carrier does not know the forwarder’s type prior to offering the menu of contracts. The carrier’s (unconditional) expected profit is obtained by taking the expectation with respect to the prior distribution.

Inspection of the proof of Theorem 1 also reveals that in the setting of the corollary, the carrier should offer contracts with a refund rate $r_i \equiv r$ that is “just below” the forwarder’s margin $p$. More precisely, the supremum in Problem $A$ is not a maximum, but there exists a sequence of feasible solutions with objective values that approach the supremum and associated informational rents that decrease to zero. These solutions can be obtained by letting $r \uparrow p$ as in the final paragraph of the proof of the theorem. The supremum is not a maximum because setting $r$ exactly equal to $p$ would cause the forwarder to book nothing and return the entire allotment for a refund.

It is also worth mentioning that it is unlikely that a carrier would in practice enter into a contract that would allow the forwarder to earn money by simply doing nothing and returning its allotment to obtain a refund. In the results above, no such contracts are offered. Nevertheless, the need to avoid such a unfavorable (to the carrier) outcome is one factor that may prevent the attainment of a first-best solution under asymmetric information when the margin is private information. If (say) $p_i < p_j$ for some types $i$ and $j$, then any contract with a refund rate $r_j$ just below the margin $p_j$ intended for a type $j$ forwarder will allow a type $i$ forwarder to select that contract and make money from the carrier by booking nothing and returning the allotment for a refund.
To close this section, we outline why two-parameter contracts (with \( r_i \equiv 0 \)) cannot eliminate informational rent or attain the first best in Problem A under the conditions of Corollary 1, even if we assume that the forwarder returns its unused capacity in advance of departure [that is, even if we assume that (1) and (6) are valid with \( r = 0 \)]. Suppose that \( n = 2 \). If we are to have \( \xi = \theta_1 \zeta_1 + \theta_2 \zeta_2 \), then we need that \( x_i = x_i^* \), and the IR constraints must be tight so that \( f_i = v_i(x_i^*) - \epsilon \) for \( i = 1, 2 \). Substituting for \( f_1 \) and \( f_2 \) in the IC constraint (15) for \((i, j) = (2, 1)\) and rearranging, we obtain \( v_1(x_1^*) \geq v_2(x_1^*) \). The stochastic monotonicity of the demand distributions in Assumption A1(i) implies that \( v_2(x) \geq v_1(x) \) for any allotment \( x \), and hence the IC constraint generally will not hold. Intuitively, the carrier cannot prevent a type-2 forwarder from choosing the contract designed for a type-1 forwarder.

4 Concluding Remarks

The International Air Transportation Association (IATA) compiles yearly industry statistics. In its September 2008 update, IATA reports that the 2007 worldwide revenue from air cargo operations was $54 billion, or 14% of the $384 billion passenger revenue (IATA 2008). Clearly, cargo operations are important for airlines (carriers). Carriers sell a significant portion of their capacity via forwarders, who act as wholesalers and provide value-adding services that the carriers are usually not equipped to perform on their own. Carriers also sell directly to some customers. In this environment, it is important for carriers to identify contract structures that lead to an overall optimal capacity allocation to the two demand streams. This is not an easy problem because forwarders have private information, which may include their demand distributions, margins, operating costs to the carrier, and reservation profits.

In this paper, we showed that if the carrier uses three-parameter contracts, where the parameters are (1) an allotment, (2) a lump-sum payment, and (3) a refund rate for unused capacity, then under certain conditions it can achieve the first-best profit. An important feature of our approach is that it obtains an upper bound on the expected informational rent paid by the carrier for a given menu of allotments, so long as certain characteristics of the forwarder that are its private information, are monotone in the forwarder type. This paper contributes to the OM literature by applying the principal-agent framework to the design
of capacity contracts in the air-cargo industry and by showing how carriers can structure capacity contracts that achieve the first-best profit.

There are several directions for future investigations within this line of research. For example, it would be interesting to consider multiple carriers and forwarders. One difficulty likely to arise in such a setting is the need to represent the dynamics of the interactions among carriers and forwarders. Even with a single carrier and multiple forwarders, it would likely be the case that the carrier would not offer contracts to all forwarders simultaneously. Developing a model that accounts for sequential or dynamic interactions between a carrier and forwarders appears to be a challenging area for exploration.

For settings with a single carrier and a single forwarder, interesting questions remain as well. For instance, one possible extension to our work is to consider settings in which the carrier has a network of flights in which shipments may move from origin to destination over multiple different routes. It may also be interesting to analyze incentive contracts that encourage forwarders to lower carriers’ service costs by mixing cargo in such a way that it fully utilizes the available weight and volume of each unit load device and by promptly picking up their loads at the destination. Since the forwarder’s effort is not observed, contracts need to overcome the underlying moral hazard problem. Another direction would be to consider situations in which some of the carrier’s demand materializes before the return of unused capacity by the forwarder. We do not allow either of the two players to overbook. However in practice, both may overbook and use a variety of rules to prioritize (reroute) loads when demand exceeds total supply. Forwarders may also try to purchase capacity in the spot market. A detailed investigation of overbooking strategies with both contractual and spot-market capacity purchases and time-differentiated bookings constitutes another area worthy of future attention.

Acknowledgments

This material is based upon work supported, in part, by the National Science Foundation under Grant No. DMI-0450359. The authors are grateful to the referees and the senior editor for their constructive comments on an earlier version of this paper.
Appendix

Proof of Lemma 1. The expression (11) follows from the first order conditions, and uniqueness follows from the strict monotonicity assumptions on $G_i(\cdot)$ and $H_i(\cdot)$.

The first step in the proof of (12) is to show the following facts.

1. If $x_i = 0$ and $\epsilon_i > 0$ then there is no way to satisfy the IR constraint (9). If $x_i = \epsilon_i = 0$, then (9) is satisfied only with $f_i = 0$, in which case $\zeta_i(x_i) = \psi_i(0)$.

2. If $0 < p_i x_i \leq \epsilon_i$, then $\zeta_i(x_i) = \psi_i(0) - \epsilon_i$.

3. If $p_i x_i > \epsilon_i$, then $\zeta_i(x_i) = \max\{\psi_i(0), \psi_i(x_i)\} - \epsilon_i$.

Fact 1 is readily seen to be true.

For fact 2, it follows from our assumptions on $G_i(\cdot)$ that $z_i(x_i) > 0$ because $x_i > 0$. Therefore, $p_i x_i - (p_i - r_i) z_i(x_i) - f_i < p_i x_i - f_i$ for any $r_i \in (0, p_i)$. Hence, $p_i x_i - (p_i - r_i) z_i(x_i) - f_i < \epsilon_i$ for all $(r_i, f_i)$ such that $r_i \in (0, p_i)$ and $f_i \geq 0$, which shows that we cannot satisfy the constraint (9) with $r_i \in (0, p_i)$. Hence, we must take $r_i \geq p_i$, in which case we can get a feasible solution by taking $r_i$ large enough. Upon choosing $f_i$ to make the IR constraint (9) tight, we obtain $\zeta_i(x_i) = w_i(0) - \epsilon_i = \psi_i(0) - \epsilon_i$.

For fact 3, if $r_i \in (0, p_i)$ then $\rho_i(x_i, f_i, r_i) = \psi_i(x_i) - \pi_i(x_i, f_i, r_i)$. Hence, the objective function in (8)–(9) is decreasing in $\pi_i(x_i, f_i, r_i)$, and therefore it is best to make the IR constraint (9) tight. When $p_i x_i > \epsilon_i$, we may choose $r_i$ close enough to (but strictly less than) $p_i$ so that this can be done, giving $\rho_i(x_i, f_i, r_i) = \psi_i(x_i) - \epsilon_i$. Similarly, if $r_i \geq p_i$, then (9) can be made tight by choosing $f_i$ appropriately, giving $\rho_i(x_i, f_i, r_i) = \psi_i(0) - \epsilon_i$.

Therefore, $\zeta_i(x_i) = \max\{\psi_i(0), \psi_i(x_i)\} - \epsilon_i$ when $p_i x_i > \epsilon_i$.

Note that $p_i x_i \leq \epsilon_i$ for all $x_i \in [0, x_i^0]$ and $p_i x_i > \epsilon_i$ for all $x_i \in (x_i^0, \infty)$. From facts 1–3, we have that

$$\zeta_i = \max \left\{ \sup_{x_i \in [0, x_i^0]} \zeta_i(x_i), \sup_{x_i \in (x_i^0, \kappa]} \zeta_i(x_i) \right\} = \max \left\{ \psi_i(0), \sup_{x_i \in (x_i^0, \kappa]} \psi_i(x_i) \right\} - \epsilon_i. \quad (18)$$

The three expressions in (12) follow from (18) by considering each case separately. For instance, to understand the expression in the case that $p_i x_i^* \leq \epsilon_i < p_i \kappa$, observe that $x_i^* \leq
\[ x_i^o < \kappa \text{ when } p_i x_i^* \leq \epsilon_i < p_i \kappa. \]

It then follows from the concavity of \( \psi_i(\cdot) \) that \( \psi_i(x_i) \leq \psi_i(x_i^o) \) for all \( x_i \in (x_i^o, \kappa] \), and consequently \( \sup_{x_i \in (x_i^o, \kappa]} \psi_i(x_i) \leq \psi_i(x_i^o) \). On the other hand, \( \sup_{x_i \in (x_i^o, \kappa]} \psi_i(x_i) = \lim_{x_i \uparrow x_i^o} \psi_i(x_i) = \psi_i(x_i^o) \), where the equality follows from the continuity of \( \psi_i(\cdot) \). Hence, \( \sup_{x_i \in (x_i^o, \kappa]} \psi_i(x_i) = \psi_i(x_i^o) \), and \( \zeta_i = \max\{\psi_i(0), \psi_i(x_i^o)\} - \epsilon_i \). The cases with \( p_i x_i^* > \epsilon_i \) and \( p_i \kappa \leq \epsilon_i \) follow by similar arguments. This completes the proof of (12).

Next we turn to the statements about the first best. Suppose initially that \( \psi_i(0) \geq \psi_i(x_i^* - \epsilon_i) \), from which it follows that \( \psi_i(0) \geq \zeta_i \) by (12). Hence, the carrier and the forwarder do not do business. Consequently, the profit of the carrier is \( \psi_i(0) = \max\{\psi_i(x_i^*) - \epsilon_i, \psi_i(0)\} \) and the profit of the forwarder is \( \epsilon_i \). Therefore, the first best is attainable if \( \psi_i(0) + \epsilon_i \geq \psi_i(x_i^*) \).

Suppose now that \( \psi_i(0) < \psi_i(x_i^*) - \epsilon_i \). This condition can be equivalently stated as \( \epsilon_i < v_i(x_i^*) + w_i(x_i^*) - \psi_i(0) \). Because \( w_i(x_i^*) \leq w_i(0) = \psi_i(0) \), this implies that \( \epsilon_i < v_i(x_i^*) \). From this, it follows that \( \epsilon_i < p_i x_i^* \). Hence, \( \zeta_i = \psi_i(x_i^*) - \epsilon_i \) by (12). Together with the supposition that \( \psi_i(0) < \psi_i(x_i^*) - \epsilon_i \), the preceding implies that \( \psi_i(0) < \zeta_i \). This means that the carrier and the forwarder do not do business, and the carrier earns \( \zeta_i = \psi_i(x_i^*) - \epsilon_i = \max\{\psi_i(x_i^*) - \epsilon_i, \psi_i(0)\} \).

At optimality the IR constraint (9) will be tight, and hence the forwarder earns \( \epsilon_i \). Therefore, the first best is attainable if \( \psi_i(0) < \psi_i(x_i^*) - \epsilon_i \).

Proof of Theorem 1. Let \( \{\delta_i \geq 0 : i \in N\} \) be slack variables corresponding to the IR constraints (16). In the following, we consider \( \{r_i : i \in N\} \) such that \( r_i \in (0, p_i) \) for all \( i \in N \). For such \( \{r_i\} \) the constraints (15) and (16) can be expressed, using (3), as

\[
\begin{align*}
p_i \hat{x}_i - (p_i - r_i)z_i(\hat{x}_i) - f_i &= \epsilon_i + \delta_i \quad (19) \\
p_i \hat{x}_i - (p_i - r_i)z_i(\hat{x}_i) - f_i &\geq p_i \hat{x}_j - (p_i - r_j)z_i(\hat{x}_j) - f_j \quad \text{for } i < j \quad (20) \\
p_j \hat{x}_j - (p_j - r_j)z_j(\hat{x}_j) - f_j &\geq p_j \hat{x}_i - (p_j - r_i)z_j(\hat{x}_i) - f_i \quad \text{for } i < j \quad (21)
\end{align*}
\]

Inequalities (20) and (21) are sometimes called upward and downward IC constraints, respectively. The objective function in (14) can be written as

\[
\sum_{i=1}^{n} \theta_i p_i(\hat{x}_i, f_i, r_i) = \sum_{i=1}^{n} \theta_i [\psi_i(\hat{x}_i) - \pi_i(\hat{x}_i, f_i, r_i)] = \sum_{i=1}^{n} \theta_i [\psi_i(\hat{x}_i) - \epsilon_i - \delta_i]
\]
From (19) we have \([p_i \hat{x}_i - (p_i - r_i)z_i(\hat{x}_i) - f_i] - [p_j \hat{x}_j - (p_j - r_j)z_j(\hat{x}_j) - f_j] = \delta_i - \delta_j + \epsilon_i - \epsilon_j\). Substituting this into the left side of (20) and (21), we see that (20) and (21) can be equivalently expressed as

\[
(p_j - p_i) \hat{x}_i + (p_i - r_i)z_i(\hat{x}_i) - (p_j - r_i)z_j(\hat{x}_i)
\leq \delta_j - \delta_i + \epsilon_j - \epsilon_i \leq (p_j - p_i) \hat{x}_j + (p_i - r_j)z_i(\hat{x}_j) - (p_j - r_j)z_j(\hat{x}_j) \quad \text{for } i < j.
\]

Also from (19) we have \(f_i = p_i \hat{x}_i - \epsilon_i - (p_i - r_i)z_i(\hat{x}_i) - \delta_i\), and therefore the requirement that \(f_i \geq 0\) can be equivalently expressed as

\[
\delta_i \leq p_i \hat{x}_i - \epsilon_i - (p_i - r_i)z_i(\hat{x}_i).
\]

From the above, it follows that \(\xi(\hat{x}) \geq \tilde{\xi}(\hat{x})\) where

\[
\tilde{\xi}(\hat{x}) \equiv \sup_{\{\delta_i, r_i\}} \left\{ \sum_{i=1}^{n} \theta_i [\psi_i(\hat{x}_i) - \epsilon_i - \delta_i] : (22), (23) \text{ and } \delta_i \geq 0, r_i \in (0, p_i) \text{ for } i \in \mathcal{N} \right\}.
\]

We now explain how to solve the preceding optimization problem. Note first that the objective function (the carrier’s expected profit) is decreasing in \(\delta_i\). Motivated by this observation, we will consider \(\{\delta_i\}\) so that the lower bound in (22) will be tight for \(j = i + 1\) and \(i = 1, \ldots, n - 1\). This yields \(\delta_1 = 0\) and

\[
\delta_i = \sum_{k=1}^{i-1} [(p_{k+1} - p_k) \hat{x}_k + (p_k - r_k)z_k(\hat{x}_k) - (p_{k+1} - r_k)z_{k+1}(\hat{x}_k)] + \epsilon_1 - \epsilon_i \quad \text{for } i \geq 2.
\]

Note that \(\delta_i \geq 0\). We next verify that \(\{\delta_i\}\) defined by (24) do satisfy (22) for all \(i < j\) when \(\{r_i \in (0, p_1)\}\) are chosen suitably. To do so, we substitute (24) into (22) and use (4) to obtain, after some algebraic manipulations, the following:

\[
(p_j - p_i) \int_{0}^{\hat{x}_i} G_j(u)du + (p_i - r_i) \int_{0}^{\hat{x}_i} [G_j(u) - G_i(u)]du
\leq \sum_{k=1}^{j-1} \left[ (p_{k+1} - p_k) \int_{0}^{\hat{x}_k} G_{k+1}(u)du + (p_k - r_k) \int_{0}^{\hat{x}_k} [G_{k+1}(u) - G_k(u)]du \right]
\leq (p_j - p_i) \int_{0}^{\hat{x}_j} G_i(u)du + (p_j - r_j) \int_{0}^{\hat{x}_j} [\tilde{G}_j(u) - G_i(u)]du.
\]
Lemma 2 below shows that if \( \{r_i \in (0, p_1)\} \) satisfy \( p_i - r_i \leq p_j - r_j \) for all \( i < j \), then the inequalities (25) — and hence (22) — do indeed hold. Note that \( p_i - r_i \leq p_j - r_j \) is trivially true when \( r_i \equiv r \in (0, p_1) \) for all \( i \in \mathcal{N} \), and moreover,

\[
\delta_i = \sum_{k=1}^{i-1} \left[ (p_{k+1} - p_k) \int_0^{\hat{x}_k} \tilde{G}_{k+1}(u) du + (p_k - r) \int_0^{\hat{x}_k} [\tilde{G}_{k+1}(u) - \tilde{G}_k(u)] du \right] + \epsilon_1 - \epsilon_i = \sum_{k=1}^{i-1} \left[ (p_{k+1} - p_k) \int_0^{\hat{x}_k} \tilde{G}_{k+1}(u) du + (p_k - r) \int_0^{\hat{x}_k} [\tilde{G}_{k+1}(u) - \tilde{G}_k(u)] du \right] + \epsilon_1 - \epsilon_i
\]

(26)

The first inequality above comes from the fact that for \( k \in [1, i - 1], \hat{x}_k \leq \hat{x}_i, r < p_k \leq p_{k+1}, \) and \( \tilde{G}_k(u) \leq \tilde{G}_{k+1}(u) \) for every \( u \). The second and third inequalities hold because \( \hat{x}_1 \leq \hat{x}_i \) and \( r \in (0, p_1) \). By Assumption A2, the interval \( I \equiv ([p_1 - (p_1 \hat{x}_1 - \epsilon_1) / z_1(\hat{x}_1)]^+, p_1) \) is non-empty. Moreover, for \( r \in I \), the term inside square braces in (27) is non-negative. Hence, for \( r \in I \), we have \( \delta_i \leq p_i \hat{x}_i - (p_i - r) z_i(\hat{x}_i) - \epsilon_i \) so that (23) holds.

To summarize, when \( r_i \equiv r \in I \) for all \( i \in \mathcal{N} \), then \( \delta_i \) in (24) simplifies to (26). Moreover, (22) and (23) hold. Therefore, \( \xi(\hat{x}) \geq \tilde{\xi}(\hat{x}) \geq \sum_{i=1}^{n} \theta_i [\psi_i(\hat{x}_i) - \epsilon_i - \delta_i] \) where \( \{\delta_i\} \) are given by (26) and \( r_i \equiv r \in I \) for all \( i \in \mathcal{N} \). Letting \( r \uparrow p_1 \) completes the proof.

\#

Lemma 2. Suppose Assumptions A1(i), A1(ii), and A2(ii) hold, and that \( \{r_i \in (0, p_1)\} \) are such that \( p_i - r_i \leq p_j - r_j \) for all \( i \leq j \). Then inequality (25) holds.
Proof. Consider \( i < j \). It follows from Assumption A1(i) and \( \hat{x}_1 \leq \hat{x}_2 \leq \cdots \leq \hat{x}_n \) that

\[
(p_i - r_i) \int_0^{\hat{x}_i} (\bar{G}_j(u) - \bar{G}_i(u)) \, du = (p_i - r_i) \sum_{k=i}^{j-1} \int_0^{\hat{x}_k} (\bar{G}_{k+1}(u) - \bar{G}_k(u)) \, du
\]

\[
\leq \sum_{k=i}^{j-1} (p_k - r_k) \int_0^{\hat{x}_k} (\bar{G}_{k+1}(u) - \bar{G}_k(u)) \, du
\]

\[
\leq (p_j - r_j) \sum_{k=i}^{j-1} \int_0^{\hat{x}_j} (\bar{G}_{k+1}(u) - \bar{G}_k(u)) \, du
\]

\[
= (p_j - r_j) \int_0^{\hat{x}_j} (\bar{G}_j(u) - \bar{G}_i(u)) \, du
\]

Therefore, (25) holds so long as (17) in Assumption A2(ii) remains valid. \( \# \)

Proof of Corollary 1. We have \( \xi \leq \sum_{i=1}^n \theta_i \zeta_i \). To prove the reverse inequality, we apply Theorem 1. From the assumptions of the corollary, Assumptions A1 and A2 are easily seen to hold. Moreover, Assumption A3(i) that \( D_0 \) and \( D \) are conditionally independent given \( T \) implies that \( D_0 \) and \( D \) are conditionally independent given \( T = i \) for all \( i \in \mathcal{N} \). Hence, \( \psi_i(x) \) is maximized by \( x_i^* \) given by (11) for each \( i \in \mathcal{N} \), and \( \mathbf{x}^* = (x_1^*, \ldots, x_n^*) \) satisfies \( x_1^* \leq x_2^* \leq \cdots \leq x_n^* \) by Assumption A3(ii). Thus, we may apply the theorem with \( \hat{x} = \mathbf{x}^* \).

We have \( \hat{\delta}_i = 0 \) for all \( i \in \mathcal{N} \) because \( p_1 = \cdots = p_n \) and \( \epsilon_1 = \cdots = \epsilon_n \). Therefore,

\[
\xi \geq \xi(\mathbf{x}^*) \geq \sum_{i=1}^n \theta_i [\psi_i(x_i^*) - \epsilon_i] = \sum_{i=1}^n \theta_i \zeta_i
\]

by Theorem 1, where the final equality follows from (12) in Lemma 1 and the fact that \( px_i^* \geq px_i^* > \epsilon \) for all \( i \in \mathcal{N} \). Therefore, \( \xi = \sum_{i=1}^n \theta_i \zeta_i \).

When \( \psi_i(0) < \psi_i(x_i^*) - \epsilon \) for all \( i \in \mathcal{N} \), we have \( \zeta_i = \psi_i(x_i^*) - \epsilon > \psi_i(0) \) for all \( i \in \mathcal{N} \) and \( \xi = \sum_{i=1}^n \theta_i [\psi_i(x_i^*) - \epsilon] = \sum_{i=1}^n \theta_i \max(\psi_i(x_i^*) - \epsilon, \psi_i(0)) \geq \sum_{i=1}^n \theta_i \psi_i(0) \).

\( \# \)

References


M. Bazaraa, J. Hurley, E. Johnson, G. Nemhauser, J. Sokol, E. Chew, C. Huang, I. Mok,


