The Contract Transformation:
A Framework for Analysis of Network Supply Chains

Pranab Majumder
majumder@duke.edu
Fuqua School of Business
Duke University
Durham, NC 27708

Ashok Srinivasan
ashoksri@usc.edu
Marshall School of Business
University of Southern California
Los Angeles, CA 90089

October 2005

Abstract

We propose a framework under which we can analyze contracting in large decentralized supply chain networks with multiple retailers, suppliers and intermediaries. Our model of a supply chain is an acyclic network where one member is designated as the contract initiator, and the sequence of contracts progresses through the tiers of the supply chain. Any member of the supply chain, manufacturer, intermediary, or retailer, can be the leader. We formulate a transformation which allows us to reduce large-scale supply chain networks to one reduced problem for the leader. The presence of a common supplier upstream of the leader reduces the contracting power of the leader, and in some cases may eliminate the “leader” from the equilibrium solution. This framework may be used to extend existing supply chain and distribution research so that multiple tiers, combined assembly and distribution, and different positions of the leader can be studied.

Keywords: Stackelberg Game, Double Marginalization, Contracting, Non-Cooperative Games, Supply Chains, Coordination
1 Introduction

Supply chains today are not limited to one or two tiers, but are in fact large interconnected and decentralized networks. For example, major auto manufacturers procure parts and subassemblies from suppliers in multiple countries, have assembly locations in multiple continents, and sell in many locations within each country. Large supply chains are common in the apparel industry as well. Other examples of large supply chains are the *keiretsu* in Japan, and the *chaebol* in Korea. These are business conglomerates spanning the complete supply chain in a virtually integrated form. They involve long-term relationships between buyers and suppliers that work together to make the supply chain more effective and efficient.

We propose a framework under which we can analyze large decentralized supply chain networks. A critical aspect of these supply chains is how each member relates to other member in terms of contracts. These contracts can range from simple, arms-length wholesale price contracts, to very complex, quasi-vertical-integrated relationships. In this paper we focus on wholesale-price and two-part-tariff contracts. An important aspect of contracting is who offers the contract to whom. We examine the entire contracting process starting with one member (the contract initiator, or leader) initiating the contract, which is followed by the contracts propagating through the tiers of the supply chain from neighbor to neighbor, both upstream and downstream. In this framework, any member of the supply chain may potentially be the leader.

Large retailers are often just the visible part of a large and complex supply chain that may start from design of the products to manufacturing, fulfillment and final sales to the consumers. The retail furniture chain Ikea designs its products in a unique manner in conjunction with its suppliers in order to determine selling price points (Margonelli, 2002). Another ubiquitous example of a retailer as the leader of a supply chain is the example of Wal-Mart, whose increasing importance to suppliers is amply demonstrated by the number of supplier liaison offices in Bentonville, AK (Hays, 2002). There are many examples of manufacturers as the leaders- semiconductor manufacturers like Intel, pharmaceutical companies like Eli Lilly, etc.

Examples abound of intermediary members (i.e. those who procure components from suppliers, and also depend upon other parties to sell the goods) who are the dominant members of the chain. Consumer goods companies like Tommy Hilfiger and Sara Lee in apparel (Anon, 1998; Shoulberg, 1997), Nike in footwear (Austin and Aguilar, 1988; Rosenzweig, 1994) and others source from many low cost (and local or global) suppliers, and reach the consumer through different retailers. In the consumer electronics industry, Ericsson in cell phones (Thurm, 2001) and Palm in personal digital assistants outsource manufacturing of the product to contract manufacturers like Flextronics (Huckman and Pisano, 2004), and sell through electronics retailer including Best Buy, Radio Shack and Office Depot.

One area of research application of the framework developed in this paper is the emergence of retailer power and its consequences on consumer welfare and channel profits. There has been a
renewed interest in the theoretical and empirical investigation of retailer power, which is also framed in terms of countervailing power (Chen, 2003; Dobson and Waterson, 1999, 1997; Ungern-Sternberg, 1996). We can model the shift in power from one member to another and examine the consequences of change of leadership on different members of a supply chain.

Another application is to the area of multinational supply chains (Johansson and Yip, 1994; Yip, 1992, 2003). The focus of this research includes multinational procurement, production and retailing, and considers the strategic aspects of working in a global context. We can apply our model to multinational supply chains which procure from multiple countries, manufacture and assemble in different locations, and sell globally in different markets. Business decisions for such large-scale supply chains thus depend upon a number of external and internal factors, for example, exchange rate changes, the extent of competition in particular markets, the efficiency of the distribution channel and economies of scale.

A third application is the study of tightly knit supply chains like Japanese keiretsu or Korean chaebol in order to examine issues of supply chain efficiency and competition between two or more such networks. A number of authors have studied these supply networks (Ahmadjian and Lincoln, 2001; Asanuma, 1993), as well as their influence on US supply chains (Dyer, 1996; Dyer and Ouchi, 1993; Ferguson, 1990; Martin et al., 1995). We can also model competition between two such networks. For example, when two auto manufacturers compete, it is not just the manufacturers, but the entire family of suppliers and distributors who compete against each other.

Finally, in addition to analysis of large scale supply networks, the framework presented in this paper can also form a basis for evaluation of alternate supply network designs. Outsourcing is no longer a choice among manufacturing locations; it is often a choice of intermediaries who in turn design the rest of the supply chain in order to deliver value. An excellent example of this is the firm of Li & Fung, who are literally supply chain architects, and manage a complex and dynamic set of suppliers, manufacturers, transporters and other partners. “Li & Fung works with an ever expanding network of thousands of suppliers around the globe, sourcing clothing and other consumer goods ... (and) draw(s) on ... expertise in distribution process technology...” (Magretta, 1998).

The contribution of this work is as follows (1) we propose a modelling framework for analysis of contracting in large decentralized supply chain networks, (2) we formulate a transformation which allows us to reduce large-scale supply chain networks into one reduced problem for the leader, and (3) we show that the presence of a common supplier upstream of the leader often reduces the power the leader has over the supply chain. For networks where the leader is upstream of all common suppliers, if the leader buys/sells nothing, then the entire supply chain buys and sells nothing. In presence of a common supplier upstream of the leader, it is no longer guaranteed that the leader will be a part of the non-zero equilibrium solution (i.e. he may buy/sell a zero quantity even when others buy/sell positive quantities), effectively reducing the contracting power of the leader.

When the supply chain leader is downstream of an upstream common supplier, it introduces considerable complexity in the problem analysis. We consider two classes of networks. In the first
class of networks, any common suppliers selling to more than one member of the supply chain are all downstream of the leader. In the second class of networks, there is at least one common supplier who is upstream of the leader.

The rest of this paper is organized as follows. After the literature review in Section 2, Section 3 presents the network model of a supply chain selling in multiple markets. We first define the network, then define the member objective functions, and end by specifying the complete game. Section 4 presents the wholesale-price contract transformation which is the central tool we use to analyze these network models. In Section 5 we discuss the solution procedure for networks without any common suppliers upstream of the leader, and in Section 6 we do the same for networks where there is a common supplier upstream of the leader. In Section 7 we present the two-part tariff contract transformation. Finally, Section 8 concludes the paper. Proofs of all lemmas and propositions are in appendix 1. Finally, appendix 2 contains the details of the two-part tariff section.

2 Literature Review

We take an industrial organization (IO) perspective on large-scale supply networks. A majority of the existing models in the field of supply chain (and the related field of channel coordination) have considered two-tier supply chains, and our work extends the analysis of supply chains to larger structures. Specifically, our work is a generalization of the von Stackelberg formulation in Spengler (1950) to model acyclic supply networks where any member of the chain can be the overall Stackelberg leader.

The literature in this area comes from three main sources.

Existing work in the Operations Management area has considered two-stage supply chains (see Tayur et al. (1999) for numerous references from this field; interesting research along such lines includes Ha (2001), Ferguson et al. (2001), Erkoc and Wu (2002) and Corbett and Tang (1999).). However, very few authors have considered larger networks. Corbett and Karmarkar (2001) also consider a multi-tier supply chain which is essentially linear in nature and has Cournot competition at each tier, examining the entry and exit of members at each stage of the supply chain; additionally, they study the impact of fixed and variable costs on competition at each stage and vertical integration. They do not study contracting; their focus is on competition at each tier and the effect of the number of members in each tier on the total quantity sold through the chain and prices at each level. Carr and Karmarkar (2003) extend the previous work into an assembly-type network. One example of a study considering larger supply chain models is by Bernstein and DeCroix (2003), who analyze an assembly network. In all these examples the supply chain is assembly-type and does not have distribution characteristics. Finally, this paper is a generalization of the linear supply chain from Majumder and Srinivasan (2005) to a network structure with multiple retailers (or selling locations).

In the Marketing literature, early channel coordination work (Jeuland and Shugan, 1983; Moorthy, 1987) study simpler manufacturer-retailer type supply chains, and consider channel coordination
questions for them. Channel coordination on longer channels has not received as much attention, and our framework is a first step towards addressing this. Similarly, research on the benefits of exclusive retailers versus common retailers (Choi, 1991) uses a model with two manufacturers and one common retailer. Other research considers industry structure and vertical integration in context of two-member supply chains (see, for example, McGuire and Staelin (1983, 1986)). Coughlan and Wernerfelt (1989) consider sequential contracting between multiple levels of a serial supply chain and coordination using 2-part tariffs. Our model can serve as a framework to examine similar research issues for larger network supply chains.

Recent work in the Industrial Organization and Economics literature (Kranton and Minehart, 2000a,b, 2001) considers bargaining between two tiers—sellers and buyers. Our work differs in that we consider multiple tiers, and that we consider a leader-follower framework. Klemperer and Meyer (1986) consider competing retailers and provide a graphical representation of price competition versus quantity competition. Our paper is similar in spirit but considers the contracting sequence in a supply chain without retailer competition. There has also been renewed interest in the concept of the ‘countervailing power’ (Galbraith, 1952) of retailers. The focus of this literature is on modelling and measuring actual retailer power (Chen, 2003) and its effect on retail prices (Dobson and Waterson, 1997). We analyze supply chains where different members (including retailers) may have contracting power.

3 A Network Model of a Supply Chain

3.1 Network Model

In this section we will construct an acyclic network model for a supply chain with multiple retailers, multiple intermediate members and multiple manufacturers. The definitions are illustrated in Figure 1.

Let a supply chain have $n$ members (or nodes) indexed from 1 to $n$. For any variable, the relevant member is denoted by a superscript, e.g., $\pi^r$ denotes the profit function for node $r$. Each supply chain is an undirected acyclic network, or a tree, in which any node is connected to any other node by a series of arcs. A member who is connected to the rest of the supply tree with only one arc is a leaf. An arc is denoted by $(j,k)$, i.e., it starts from node $j$ and ends at node $k$, and represents a relationship between the two members. The path between any two nodes (i.e. an ordered sequence of undirected arcs) is unique. Let the set of arcs in the supply chain be denoted by $T$.

$$ T = \{(j,k) : j, k \in \{1..n\}, j \neq k\}. $$

Since the supply network is an acyclic network, the number of elements in $T$ is $n - 1$. Two members are neighbors if there is an arc in $T$ which connects them. Relationship between any two neighbors are both goods-flow relationships (who supplies whom) as well as contract relationships (who offers a contract to whom).
Without loss of generality, we index the manufacturing nodes first, so that they have the lowest indices. Similarly, we index the retailer nodes last, so that no non-retailer node has an index higher than any retailer node. For convenience, all other nodes may be called intermediate nodes.

We define the distance function \( \delta(j, k) \) to be the distance between nodes \( j \) and \( k \) in \( T \). For example, \( \delta(j, j) = 0 \), and \( \delta(j, k) = 1 \) \( \forall j, k : (j, k) \in T \). Since this is an acyclic network, \( \delta(j, k) \) is unique for any \( j, k \).

In order to denote the flow of goods, we consider directed arcs starting from manufacturing nodes to the intermediaries, and the forward path eventually ending with some retailer node. This is defined to be the “supply tree”. Each manufacturing or intermediary node has at least one arc originating from it. Each intermediate or retailer node has at least one arc ending on it. Manufacturing nodes do not have any arc ending on them, and retailer nodes do not have any arc starting from them. Intermediary nodes have arcs starting from them as well as arcs ending on them. For a supply network, the supply tree is a set \( S \) of \( n - 1 \) directed arcs. The sets \( T \) and \( S \) have the same arcs; the only difference is that the arcs in \( S \) are directed, while in \( T \) they are undirected.

\[
S = \{(j, k) : (j, k) \in T \text{ or } (k, j) \in T, k > j \}.
\]

If \( (j, k) \in S \), we will say that \( j \) is upstream of \( k \), and \( k \) is downstream of \( j \). Thus manufacturing leaves are downstream of no other node, the retailer leaves are upstream of no other node, and intermediary nodes are both upstream of some node(s) and downstream of some node(s). All intermediary

Figure 1: The contract tree and the network definitions.
nodes are numbered and connected in such a fashion that the arcs always point from a lower index node to a higher index node.

Define the set of manufacturing nodes and the set of retailer nodes as

\[
M = \{ j : j \in \{1..n\}, \exists k \text{ s.t. } (k,j) \in S \}
\]

\[
R = \{ j : j \in \{1..n\}, \exists k \text{ s.t. } (j,k) \in S \}.
\]

In order to denote the leadership structure in the network we define member \( i \) as the leader. We will take the location of the leader as exogenous, rather than incorporate that as a decision in these models. For an analysis of optimal leader location in linear supply chains, see Majumder and Srinivasan (2005). The leader offers contracts to her neighbors in the network (as defined in the arc set \( T \)). Subsequently each member who received a contract from the leader offers a contract to all his neighbors (except the leader) and this process continues until all members have been offered a contract. This is defined to be the “contract tree”. The contract tree is a set \( C \) of directed arcs. The sets \( T \) and \( C \) have the same arcs; the only difference is that the arcs in \( C \) are directed, while in \( T \) they are undirected.

\[
C = \{ (j,k) : (j,k) \in T \text{ or } (k,j) \in T \}.
\]

If \((j, k) \in C\), we will say that \( j \) is “parent” of \( k \), and \( k \) is “child” of \( j \). Figure 1 shows an example of a contract tree.

Thus, \( T \) defines an acyclic network (tree) with undirected arcs, \( S \) defines a tree with directed arcs all pointing towards the retailers, and \( C \) defines a tree with directed arcs all pointing away from the leader \( i \). In \( S \) each node except the retailers has arcs originating from it (the retailers have none). The leader node is the root of the tree \( C \) and has no arc ending on it; every other node except the leader has exactly one arc ending at it.

We will also define a branch point, applicable to networks with more than one retailer node. A branch point is a node that sells to more than one downstream node. Thus, a network that has multiple retailer nodes must have at least one branch point. Thus, node \( j \) is a member of the set of branch point nodes \( B \) if it has more than one downstream node, or,

\[
j \in B \iff |\{ k : (j,k) \in S \}| > 1
\]

Each branch point node may supply different quantities to all its the downstream nodes. For a branch node \( m \), we will use the notation \( q^{m,n} \) to denote the quantity that node \( m \) supplies node \( n \). (If a particular node is not a branch point then there is no ambiguity.)

Branch points are illustrated graphically using supply trees in Figure 2. Whether the branch point is upstream or downstream of the leader is an important factor.

We will additionally define some sets associated with some nodes. For a retailer/market \( r \) define the “supplier subset” \( S^r \) to contain all nodes such that there is a directed path from any node in \( S^r \) to \( r \). By definition, \( r \) itself is also an element of \( S^r \). The supplier-subset is defined for each retailer, but not for any other node.
Similarly, for any node $j$, define the “retailer subset” $R^j, R^j \subset R$ as the set of retailers that finally receive the goods processed at $j$, i.e., $R^j = \{k : k \in R, j \in S^k\}$. The retailer-subset is defined for any node, including the retailers themselves. If node $j$ is a retailer, i.e., if $j \in R$ then $R^j = \{j\}$.

3.2 Cost Structure and Procurement Relationships

Each node has a marginal cost function $MC^k(q)$. There are no fixed costs, i.e. all total cost functions start from zero. We distinguish between three groups of nodes. (a) The marginal cost function for any manufacturing node is strictly increasing. We will assume that $MC^k(q)$ is non-negative, non-decreasing, convex and differentiable at least twice. (b) For all other nodes except the leader, the marginal cost function may be increasing, constant, or zero. (c) The leader can have any type of marginal cost function: increasing, decreasing, constant, U-shaped, etc. (For expositional convenience we will assume that the leader also has a marginal cost function that is increasing, constant or zero. However, the methodology will work for any other cost function as well.)

The nature of marginal costs as a function of quantity is an ongoing debate, and it is clear that any one class of cost functions is not applicable for all situations. We employ increasing marginal costs, and to this end provide the arguments in its support. These arguments can be broadly classified into the following groups.

1. Congestion related: There are many reasons why a facility may have an increasing marginal cost. This may be because (a) a facility may have some defined “capacity” which can not be completely utilized in practice (for example, costs increase dramatically when production volume approaches the capacity); (b) many factors, like managerial expertise can not be scaled perfectly, leading to increasing costs. Banker et al. (1988) observe that stochasticity and variability in production environments may cause marginal costs to increase with volume due to congestion and queuing delays. The authors examine US industry-level capacity utilization data and find that job shop and batch processing industries (compared to mass production and continuous
process industries) operate at lower utilization levels in order to control these costs.

2. Existing literature: The production smoothing model described in Holt et al. (1960) and later work in the same spirit (see Akella et al. (1992); Zipkin (1982)) uses a quadratic total cost, which implies a (linear) increasing marginal cost of production. A U-shaped marginal cost is commonly described in economics and theory of the firm (see Tirole (1988) and McAuliffe (1999)), and an increasing marginal cost is a simplification. Finally, Eliashberg and Steinberg (1991) provide an excellent review of the literature supporting increasing marginal costs in the context of their model of competition between a firm with constant marginal costs and a firm with increasing marginal costs.

3. Empirical evidence: In addition to Banker et al. (1988) already listed, Mollick (2004) studies a range of Japanese vehicle industries from bicycles to large busses and observes that for 6 out of the 9 goods studied, the industry operates in the range of increasing marginal costs. Similarly, Griffin (1972) examined the US petroleum refining industry and obtained increasing short-run marginal costs. In the microelectronic industry, in the short run the supply curve implies an increasing marginal cost (Varian, 1987) (pg. 383-404). Finally, Blanchard (1983) and Kashyap and Wilcox (1993) find empirical evidence in the US automobile industry consistent with the production smoothing model in Holt et al. (1960) which has increasing marginal costs.

4. Contracting requirements: The purpose of our research is to examine issues pertaining to contracting. If all marginal costs are constant, it is not possible for the downstream member to offer a two-part tariff or wholesale price contract to the upstream member, since the contracted quantity will then either be zero or unbounded. Employing an increasing marginal cost function allows us to analyze contracts offered by the downstream member.

This cost function allows us to model two types of procurement relationships:

- **Market procurement:** Some components may be standardized to such an extent that they can be (relatively) easily procured from a marketplace. Some examples are tires, nuts, bolts, computer memory, hard disks etc. Since there exists a competitive market for these components, an individual firm cannot change the price at which these goods are bought and sold. Thus, these items have a constant marginal cost of procurement. For any member of the supply chain we will assume that the cost of procurement of all such components are incorporated into the cost function of the member procuring these items.

- **Relationship-specific procurement:** Procurement relations may be specific to the supplier and the buyer for customized products which cannot be sold to other buyers without incurring additional significant costs, if at all. Examples are anti-lock brake systems, application specific integrated circuits, etc. If the seller has a specific relationship with the buyer, then we model the buyer and the seller as separate members in the supply chain, and explicitly model the seller’s cost function.
3.3 Demand

For a retailer node \( r \in R \) the demand function is \( p^r(q) \), with the associated marginal revenue function

\[
MR^r(q) = \frac{\partial}{\partial q} qp^r(q).
\]

We require the demand function to be decreasing, differentiable at least twice, and to have decreasing elasticity as \( q \) increases. This is satisfied by any concave demand function, and also by a linear demand function. In addition, many convex demand functions, like \( p = ce^{-q} \) also satisfy this condition. However, the constant elasticity demand function \( p = c \frac{1}{q} \) does not satisfy this. For expositional convenience, we assume that \( p^r(q) \) is decreasing, concave and differentiable at least twice.

3.4 Member Objective Functions

We will now specify each member’s objective function. We will use the contract tree \( C \) and the supply tree \( S \) extensively in defining these objective functions. We will denote any arbitrary member as node \( s \), its parent node as node \( r \), and its child nodes as \( t_1, t_2 \), and so on. Node \( s \) is offered a wholesale price by the parent node \( r \) such that \((r, s) \in C \) and offers contracts to all nodes \( t \) such that \((s, t) \in C \). Note that for the leader there is no parent node \( r \), and for a leaf in the network \( C \) there is no child node \( t \). Let the quantity selected by member \( s \) be \( q^s \). Let the wholesale price offered to member \( s \) be \( w^s \).

3.4.1 Non-Retailer Objective Functions

We will first consider the objective function for all nodes except the retailer node. The retailer nodes have a demand function, and hence their objective functions are different. In general, the parent \( r \) may be either upstream or downstream of \( s \). Some of the child nodes \( t \) may be upstream and some may also be downstream of \( s \).

We will define some additional variables. Consider a node \( s \). First, consider whether \( s \) is offered a contract from a downstream parent, i.e., \( I_d(s) \) equals 1 only if \( s \) has a downstream parent

\[
I_d(s) = \begin{cases} 
1 \text{ if } \exists r \text{ s.t. } (r, s) \in C \text{ and } (s, r) \in S \\
0 \text{ otherwise}
\end{cases}
\]

Next, consider whether \( s \) is offered a contract from an upstream parent, i.e., \( I_u(s) \) equals 1 only if \( s \) has an upstream parent

\[
I_u(s) = \begin{cases} 
1 \text{ if } \exists r \text{ s.t. } (r, s) \in C \text{ and } (r, s) \in S \\
0 \text{ otherwise}
\end{cases}
\]
Finally, define the two sets of children- $C_d(s)$ for the downstream children, and $C_u(s)$ for the upstream children

$$C_d(s) = \{t : (s,t) \in C \text{ and } (s,t) \in S\},$$

$$C_u(s) = \{t : (s,t) \in C \text{ and } (t,s) \in S\}.$$  

Figure 3 provides examples of these additional variables. Part (a) shows the indicator variables and children sets for a node which has an upstream parent, part (b) shows these for a node which does not have any parent (i.e., the leader), and part (c) shows the variables for a node with a downstream parent.

The node $s$ is offered a wholesale price $w^s$ by the parent node $r$. (Note that if member $s$ is the leader of the contract tree, then there is no member who offers a contract to $s$, and hence there is no such $w^s$.). Node $s$ must decide the quantity to produce/process $q^s$ as well as the wholesale prices $w^t$ to offer to all children nodes $t$. Thus, the profit function for any member $s \in \{1..n\} \setminus R$ (i.e., except
a retailer) is
\[
\max \pi^s \left( \{q^{s,t}\}_{t \in C_d(s)}, \{q^{s,r}\}_{(r,s) \in C}, \{w^t\}_{(s,t) \in C}, \{w^s\}_{(r,s) \in C} \right) \tag{6}
\]
\[
= w^s q^{s,r} I_d(s) - w^s q^s I_u(s) + \sum_{t \in C_u(s)} u^t q^t - \sum_{t \in C_u(s)} w^t q^{t,s} - \int_{x=0}^{q^s} MC^s(x) \, dx
\]
\[
\text{s.t. } q^s \leq q^{t,s} \quad \forall t \in C_u(s) \cap B, \text{ (upstream children that are branch points)}
\]
\[
q^s \leq q^t \quad \forall t \in C_u(s) \setminus B, \text{ (upstream children that are not branch points)}
\]
\[
q^{s,t} \geq q^t \quad \forall t \in C_d(s), \text{ (downstream children)}
\]
\[
\text{and } q^s \geq I_d(s) q^{s,r} + \sum_{t \in C_d(s)} q^{s,t}.
\]

The variables listed before the semi-colon in the profit function definition are the decision variables for $s$. The variables listed after the semi-colon are the parameters for node $s$. There may be at most one parameter- $w^s$, the wholesale price that $r$ has offered $s$. Note that $\{q^{s,r}\}_{(r,s) \in C}$ is a (single) decision variable only if $s$ has a downstream leader. Similarly, $\{w^t\}_{(s,t) \in C}$ is the set of wholesale prices that $s$ offers the downstream children.

In the objective function, the first term is the payment received from the downstream parent, if any. The second term is the payment made to the upstream parent, if any. The third term is the payment received from all downstream children. The fourth term is the payment made to all upstream children. Finally, the fifth term is the cost of processing the goods. Since any node has at most one parent some of these terms will be zero.

The first two constraints state that node $s$ cannot produce more than what any upstream child supplies. If the upstream child $t$ is a branch point, this constraint is $q^s \leq q^{t,s}$, otherwise it is simply $q^s \leq q^t$. The third constraint states that the quantity $q^{s,t}$ that $s$ provides to each downstream child $t$ must cover the quantity $q^t$ that that child demands (since this $q^t$ depends upon the $w^t$ that $s$ offered to begin with). The last constraint states that the quantity the branch point produces cannot be less than what it supplies to all the downstream nodes.

These constraints restrict the feasible region for $s$’s decisions. For example, consider a node $s$ with an downstream child $t$, and an upstream parent $r$. It is infeasible for $s$ to offer $t$ a wholesale price $w^t$ that will induce $t$ to select a quantity $q^{s,t}$ that is greater than the quantity that $s$ decides to procure from $r$, $q^{r,s}$. If indeed $q^{s,t} > q^{r,s}$ for some values of $w^t$ and $q^{r,s}$, $s$ can either reduce $w^t$ or increase $q^{r,s}$ to reach feasibility. From the game-theoretic perspective, this rules out illogical but theoretically possible situations like one in which $s$ offers $t$ a positive wholesale price $w^t$, but arbitrarily decides to process nothing. Later in this section we show that in equilibrium, these constraints are binding (in Corollaries 1, 3 and 5).
3.4.2 Retailer Objective Function

We next consider the objective function for a retailer. This differs from the previous one in two ways. First, there is no downstream node. Hence, there cannot be a downstream parent node, and there are no downstream children. Thus, these terms are absent in the objective function. Second, a retailer faces a demand function. This accounts for the new term in the objective function. Once again, we assume that the parent node, if any, is node \( r \), and the children nodes are nodes \( t \). The retailer herself is node \( n \). Figure 4 shows two examples of a retailer node, and lists the indicator variables and child node sets for each example.

A retailer is offered a wholesale price \( w^s \) by member \( r \). She has to decide the quantity to sell \( q^s \), and the wholesale price to offer to her children nodes \( t \). The parent node \( r \), if any, is always upstream of the retailer. If the retailer is the leader of the contract tree, then there is no such member \( r \), and hence no such \( w^s \). The retailer’s objective function is

\[
\text{max } \pi^r \left( q^r, \{w^t\}_{(r,t) \in C}; \{w^s\}_{(r,s) \in C} \right)
\]

\[
= \int_{x=0}^{q^s} MR^s(x) \, dx - w^s q^s I_u(s) - \sum_{t \in C_u(s)} w^t q^{t,s} - \int_{x=0}^{q^s} MC^s(x) \, dx
\]

s.t. \( q^s \leq q^{t,s} \forall t \in C_u(s) \cap B \),

\( q^s \leq q^t \forall t \in C_u(s) \setminus B \).

![Figure 4: Two examples of a retailer node showing the leader indicator functions and children sets.](image)

The first term in the profit function (7) represents the revenue from the market. The second term denotes the payment to the upstream parent, if any. The third term contains the payments to the upstream children (if any). Finally, the last term is the retailer’s own cost.
In the network example shown in Figure 1, node 4 is the leader. His optimization problem is

\[ \pi^4 (q^4, w^1, w^6) = w^6 q^6 - w^1 q^1 - \int_0^{q^4} MC^4 (x) \, dx \]
\[ \text{s.t. } q^4 \leq q^1, \]
\[ q^4 \geq q^6. \]

Node 6 obtains \( w^6 \) from node 4, and offers contracts to nodes 5, 7 and 8; her objective function is

\[ \pi^6 (q^6, w^5, w^7, w^8, w^6) = w^7 q^7 + w^8 q^8 - w^5 q^5 - w^6 q^6 - \int_0^{q^6} MC^6 (x) \, dx, \]
\[ \text{s.t. } q^6 \leq q^5, \]
\[ q^6 \geq q^7 + q^8. \]

Node 7 is a retailer, and her objective function is

\[ \pi^7 (q^7; w^7) = p (q^7) q^7 - w^7 q^7 - \int_0^{q^7} MC^7 (x) \, dx. \]

Since she has no children nodes, there are no constraints for her optimization problem.

### 3.4.3 Complete Game Specification

For a given supply tree, the complete game can be specified in extensive form as follows (it is not practical to explicitly write out the extensive-form game for a general tree). The game starts with the leader’s problem (node \( i \)) in the first stage as the root. The root has a number of branches, one for each of its child nodes. Each of \( i \)’s child nodes \( j \) offers contracts to its own children; thus there are sub-branches from \( j \) to each one of \( j \)’s child nodes, and so on. Thus, in each branch there is a sequence of sub-branches each consisting of the next tier members’ problems. Some of the branches progress upstream of the leader until they end at a manufacturer, and some branches progress downstream of the leader until the retailer. Some branches may first go downstream and then upstream, and vice versa.

As an example, consider Figure 5 which shows the game structure corresponding two locations of the leader from the example discussed earlier in Figure 1. Note that we obtain the game structure by pivoting the supply tree (Figure 5 (a)) around the leader node. In the first example, member 4 is the leader, and is at the root. In the second example member 7 is the leader and is at the root.

Note that the information set for any player \( j \) is the set of wholesale prices that each of her ancestors (starting from the leader or root node, and ending with her immediate parent) has offered until then. Observe that out of this, only the wholesale price that \( j \)’s immediate parent offers her, \( w^i \), is relevant to her own profits. The strategy set for \( j \) is her own quantity \( q^j \), and the wholesale prices she offers
her own children. All of these are positive variables. While the range for the wholesale prices are \([0, \infty)\), the range for \(q^j\) depends upon the quantity decisions of her immediate upstream children \(0 \leq q^j \leq \min\{q^k\}_{k \in C_u(j)}\). This range is derived from the constraints in the objective functions (6) and (7) shown earlier. Since this is a game of perfect information, each information state for \(j\) occurs at a unique decision point (see Myerson (1991)).

### 3.5 The Centralized Solution

The centralized solution to the supply tree is the solution to the problem when one member faces the demand function in each market as well as incurs all the costs at each node, and maximizes the profits for the chain as a whole, i.e.,

\[
\max_{\{q^r\}_{r \in R}} \pi^C\left(\{q^r\}_{r \in R}\right) = \sum_{r \in R} \left( \int_0^{q^r} MR^r(x) \, dx \right) - \sum_{j=1}^n \left( \int_0^{q^j} MC^j(x) \, dx \right)
\]

\[
q^r \geq 0
\]

\[
q^j = \sum_{r \in R^j} q^r.
\]

Note that the first term is the total revenue from all the retailers/markets. The second term is the total cost for all the nodes. For each node \(j\), the quantity processed \(q^j\) is related to the quantities sold by the retailers by the last equality, i.e., the node processes the quantity only for the markets belonging to its retailer set \(R^j\).

The centralized solution is obtained from the \(|R|\) first-order-conditions with respect to each \(q^r\). While the marginal revenue for market \(r\) depends only on \(q^r\), the marginal costs at any node \(j\) that
processes the units sold in market \( r \) depend upon \( q^j \), which processes units for many markets, in general. Hence we will obtain a set of \(|R|\) simultaneous equations, which depend on the particular network configuration. Each of these balance the marginal revenue from a market with the corresponding marginal costs, i.e. \( MR_r (q^r) = \sum_{j \in S^r} MC^j \left( \sum_{k \in R^j} q^k \right) \). We include the complementary slackness conditions (which state that if the network sells nothing in a particular market, then the marginal function balance for that market is not relevant). These first-order-conditions are

\[
\left[ MR_r (q^r) - \sum_{j \in S^r} MC^j \left( \sum_{k \in R^j} q^k \right) \right] q^r = 0 \quad \forall r \in R.
\]

**Lemma 1** There exists a unique solution to the centralized problem.

As we demonstrate in the appendix (Section 9.1), the objective function is concave for any general network. (All proofs of Lemmas are in the Appendix)

Before we provide the network solution, we may make some observations about the profitability of individual markets. These conditions are necessary for the network to sell in a particular market, but may not be sufficient, since the decision to sell in a particular market is not separable across markets.

**Lemma 2** For a market \( r \), a necessary condition for \( q^r > 0 \) in equilibrium is

\[
MR_r (0) > \sum_{j \in S^r} MC^j (0).
\]

This condition simply states that the network may sell in market \( r \) only if it would sell the first unit in market \( r \) in absence of all other markets (when we also eliminate the nodes not belonging to the supplier set \( S^r \)).

## 4 The Wholesale Price Contract Transformation

Before solving the general network problem we have defined in the previous section, we will illustrate the basic mechanism by which we reduce the large network problem to a much smaller problem by the process of eliminating nodes.

### 4.1 Double Marginalization Revisited

Consider the problem first described by Spengler (1950). A manufacturer with a constant marginal cost \( c \) offers a wholesale price \( w \) to a retailer who faces a demand \( p(q) = a - bq \). The retailer’s problem is

\[
\max_q \quad \pi^R (q; w) = (p(q) - w) q
\]
Similarly, the manufacturer’s problem is

$$\max_w \pi^M(w) = (w - c)q$$

given the optimal decision by the retailer for any $w$ that he sets. It is simple enough to obtain the optimal $q$ decision by the retailer as a function of $w$, and incorporate this into the manufacturer’s objective function and then solve for the optimal $w$ that will maximize the manufacturer’s profits. Spengler used this to show that this contract between the two members causes “double marginalization”, which reduces the total supply chain profits, as well as the quantity sold by the chain.

We will rework this problem to illustrate the structure of the solution. Instead of solving for the retailer’s optimal $q$, we can transform the actual demand function into an “induced” demand function for the manufacturer, and obtain the solution to the two-member supply chain by solving only the modified problem for the manufacturer. The manufacturer offers a price $w$, and as a consequence of that, sells a certain quantity. Although we know that he is actually selling to the retailer, and not the end customer, his “problem” is analogous to a seller facing a demand function. We show this in Figure 6.

If we assume that the retailer faces $p(q)$, a decreasing and concave demand function, then she maximizes $\pi^R(q) = (p(q) - w)q$, and hence chooses $q$ such that $(p(q) - w) + qp'(q) = 0$, or

$$w(q) = p(q) + qp'(q).$$

Note that $p(q) + qp'(q) = \frac{d}{dq}qp(q)$ is the retailer’s marginal revenue function $MR(q)$. Define

$$F(p(q)) = \frac{d}{dq}qp(q) = p(q) + qp'(q).$$

Figure 6: Double Marginalization in terms of the Contract Transformation
Thus, the induced demand function for the manufacturer is the actual marginal revenue function for the retailer. Now, the manufacturer solves the optimization

\[
\max_q \pi^M(q) = (MR(q) - c)q.
\]

The solution to the above problem is \(MR(q) - c + MR'(q)q = 0\), or

\[
MR(q) + qMR'(q) = c. \tag{9}
\]

Observe that \(MR(q)\) goes through the same operation \(F(\cdot)\) that we applied to \(p(q)\) to obtain \(MR(q)\) in the first place. Thus, if the retailer faces a marginal revenue function \(MR(q)\), then the manufacturer faces an induced marginal revenue function \(F(MR(q)) = MR(q) + qMR'(q)\). Thus, the two-member problem reduces to a one-member problem: a manufacturer with a constant marginal cost facing an induced marginal revenue function. In other words, we have “eliminated” the retailer node from the two-node supply chain, as shown in Figure 6 (b).

In a similar way, we can extend Spengler (1950) to consider what happens if a retailer offers a wholesale price \(w\) to a manufacturer whose marginal cost function is \(MC(q) = c + 2Kq\) (instead of a constant \(c\)). (Thus, given \(w\), the manufacturer optimizes \(\pi^M(q) = wq - \int_0^q MC(x)dx\), and the retailer optimizes \(\pi^R(w) = (p(q) - w)q\).) We can define an “induced cost function” for the retailer. If \(MC(q)\) is increasing and convex, then the manufacturer will choose \(q\) such that \(MC(q) = w\). Hence the total cost for the retailer is \(wq = (MC(q))q\). Thus the induced marginal cost function for the retailer is

\[
\frac{d}{dq}(MC(q)q) = MC(q) + qMC'(q) = F(MC(q)), \tag{10}
\]

which has a similar structure to the induced demand function in the previous example. (For \(MC(q) = c + 2Kq\), we obtain the “induced cost function” for the retailer as \(c + 4Kq\).) Thus, we have “eliminated” the manufacturer node from the two-node supply chain.

When we incorporate the optimal decision by the follower into the leader’s optimization, we are essentially solving for the Sub-Game Perfect Equilibrium of a leader-follower game. The contract transformation that we define next is a compact way of representing this folding of the follower’s optimal decision into the leader’s optimization problem. We can now generalize the operation shown in equations (9) and (10) in the next section.

### 4.2 The Transformation Definition

In the definitions which follow, we will first define the type of functions \(f\) on which the transformation \(F\) can be applied. There are two types of functions \(f\)- the first are marginal cost functions, which are increasing and convex and the second are marginal revenue functions, which are decreasing and concave.
Let $f \in C^n$ (the set of continuous $n$-times differentiable real functions mapping $\mathbb{R}$, the real line, to itself). Define the transformation $F_x(.)$ which operates on $f$ in the following manner

$$F_x(f(x)) = f(x) + x \frac{\partial f(x)}{\partial x}. \quad (11)$$

We will now list some properties of $F_x$.

**Lemma 3** $F_x(.)$ is a linear transformation, i.e., (a) $F_x(f_1 + f_2) = F_x(f_1) + F_x(f_2)$ and (b) $F_x(\alpha f) = \alpha F_x(f)$. (All proofs are in the appendix).

**Lemma 4** (a) For all positive real $x$, $F_x(f(x))$ is increasing for all increasing and convex $f$.

(b) For all positive real $x$, $F_x(f(x))$ is decreasing for all decreasing and concave $f$.

This property establishes that when we apply $F_x$ to any increasing convex marginal cost function we obtain another function which is also increasing. Similarly, when we apply $F_x$ to any decreasing concave marginal revenue function it generates another function which is also decreasing.

Define the repeated transformation in this manner: $F^n_x(f) = f$, $F^j_x(f) = F_x(f)$ and $F^n_x(f) = F_x(F^{n-1}_x(f)) \quad \forall n > 1$. We refer to these transformed functions as $N$-marginalized functions (until now we have only marginalized the functions once). In order to analyze the contract tree the marginal cost and marginal revenue functions must be $N$-times differentiable, where $N = \max_i (\delta(i,j))$, i.e., the maximum distance between the leader and any other node. In the sections that follow we establish that $F_x$ allows us to eliminate a leaf from the network without changing the quantity solution to the supply tree.

### 4.3 Decentralized Solution

In the following sections 5 and 6 we show how to use the contract transformation to successively eliminate nodes from the network by incorporating the node’s problem parameters into its parent’s problem, keeping the solution unchanged. The equilibrium concept we use is the Subgame Perfect Equilibrium, i.e., at each stage each member makes optimal choices, given the previous choices made by members in previous stages and correctly anticipating the optimal choices by members in subsequent stages (see Myerson (1991) for a detailed description of Subgame Perfect Equilibrium).

#### 4.3.1 Node Elimination Procedures

The extent to which a network can be reduced by eliminating nodes depends upon a number of factors. The most important factor is whether there are branch points which receive contracts from one of their own downstream nodes. Regardless of the presence or absence of such nodes, the rest of the network can usually be reduced by using some standard procedures. There are two standard procedures— the
first eliminates manufacturing leaves (we call this Procedure A) and the second eliminates retailer leaves (we call this Procedure B). (These procedures are shown in the next section.)

The sub-network configuration shown in Figure 7 (a) consists of a parent node $j$ and an upstream manufacturing leaf $k$. We can use Procedure A to eliminate node $k$. This is a common configuration in any supply chain, since there must be one or more suppliers.

The sub-network configuration shown in Figure 7 (b) consists of a parent node $j$ and one or more downstream retailer leaves $\{k\}_{(j,k) \in C,\{j,k\} \in S}$. We can use Procedure B to eliminate nodes $\{k\}_{(j,k) \in C,\{j,k\} \in S}$. This is also a common configuration whenever an intermediary sells to multiple retailers.

Successive use of Procedures A and B will prune the network of all manufacturing children leaves and many of the retailer leaves. For networks where the leader is upstream of all branch point nodes, these two procedures are adequate to reduce the network down to a single node (the leader). Thus, for these networks no further procedures are needed. An example of this is shown in Figure 5 (b), where the leader, node 4, is upstream of the branch point node 6.

However, we cannot use Procedure B to eliminate retailer leaves if the parent of the retailer leaves receives a contract from one of the retailers, as shown in Figure 7 (c). Examples of a configuration where the leader is downstream of the branch point can be a powerful retailer who dictates terms to a large manufacturer who also sells to other retailers (e.g. Wal-Mart offering contracts to Proctor and Gamble who sell to other retailers as well). Another example is a large car manufacturer who offers a contract to an auto parts manufacturer who also sells to other car manufacturers or the replacement parts market (e.g. General Motors purchasing from Delphi who sells to other car manufacturers and auto parts retailers).

These are networks where there are one or more branch points upstream of the leader node. An
example of this is shown in Figure 5 (c), where the leader, node 7, is downstream of the branch point node 6. Thus we cannot use Procedure B to eliminate node 8, the other retailer. If we restrict ourselves to only Procedures A and B, we will still be left with a network of three nodes- 6, 7 and 8. Depending upon the original network, this reduced network can quite large.

We will define an additional procedure C to consider such networks in Section 6. In section 5 we will restrict ourselves to networks where the leader is upstream of the first branch point. Once we have the reduced problem, we can then solve for the quantities that the network sells in each market. In addition, we also provide corollaries that allow us to calculate the wholesale prices offered throughout the network. From the member objective functions, we can also calculate the member profits.

5 Networks without Upstream Branch Points

5.1 Node Elimination Procedures and Reduction Algorithm

We will begin by providing the network solution for networks where the leader is upstream of any branch point(s). As mentioned earlier, we do this by first defining two node elimination procedures, and then by using them as components of an algorithm that reduces the network to a reduced problem for the leader.

5.1.1 Procedure A: Manufacturing Leaf Elimination

Let \((j, k) \in C\) such that \(k \in M\), and \((k, j) \in S\), i.e., \(k\) is a manufacturing leaf supplying goods to \(j\) and receiving a contract from \(j\). Member \(j\) is the downstream parent of member \(k\). This is shown in Figure 7(a). This is a local Stackelberg leader-follower game. In this procedure we will eliminate node \(k\) and incorporate her marginal functions into node \(j\)’s problem.

Their optimization problems are given by Problem 1.

**Problem 1** The optimization problem for \(j\) can be rewritten as

\[
\begin{align*}
\max \quad & \pi^j(q^j, \ldots, w^k, \ldots) \\
= \quad & w^j q^j I_d(j) - w^j q^j I_u(j) \\
& + \gamma(.) - w^k q^k - \int_{x=0}^{q^j} MC^j(x) \, dx, \\
\text{s.t.} \quad & q^j \leq q^k, \\
& \Gamma.
\end{align*}
\]

The term \(\gamma(.)\) in (12) includes \(j\)’s revenues and costs from contracting with children other than \(k\). The set \(\Gamma\) contains the other constraints from \(j\)’s optimization from contracting with children other than \(k\). Similarly, \(j\)’s objective function may have other parameters and decision variables that are not directly related to contracting with \(k\), for example \(w^j\).
The optimization problem for $k$ is

$$\max \pi^k (q^k; w^k) = w^k q^k - \int_{x=0}^{q^k} MC^k(x) \, dx. \quad (13)$$

**Lemma 5** The optimal solution to the optimization for a manufacturing leaf $k$ as shown in equation (13) is

$$q^k = \begin{cases} 
q^k : MC^k(q^k) = w^k & \text{if } w^k > MC^k(0) \\
0 & \text{otherwise}
\end{cases}$$

Lemma 5 shows that $k$’s optimal decision depends only upon $w^k$ that $j$ offers, not on the quantity that $j$ chooses. However, recollect that $j$ is constrained to choose $q^j \leq q^k$.

Define a modified problem for $j$ alone, as shown in Problem 2. Note that instead of offering a contract to $k$, $j$’s problem now includes $k$’s costs transformed by $F_{q^j}$.

**Problem 2** Rewrite the objective function for $j$ in Problem 1 by replacing $w^k q^k$ $k$’s cost function transformed by $F_{q^j}$, i.e. $\int_{x=0}^{q^j} F_x \left( MC^k(x) \right) \, dx$

$$\max \pi^j (q^j, \ldots) = w^j q^j I_d(j) - w^j q^j I_u(j) + \gamma(.) - \int_{x=0}^{q^j} F_x \left( MC^k(x) \right) \, dx - \int_{x=0}^{q^j} MC^j(x) \, dx,$$

s.t. $\Gamma$.

**Lemma 6** The optimal solution to $j$’s optimization in Problem 1 (incorporating optimal behavior by $k$) and the optimal solution to $j$’s optimization in Problem 2 have identical $q^j$ and $\pi^j$.

**Corollary 1** In Problem 1, $j$ will always set $q^j = q^k$, or, equivalently, set $w^k$ such that $k$’s optimal $q^k$ will equal $q^j$.

**Corollary 2** For an equilibrium solution where $j$ produces $q^j$ (obtained after solving the complete problem), we can calculate the wholesale price offered by $j$ to $k$ in problem 1 as

$$w^k = \begin{cases} 
MC^k(q^j) & \text{if } q^j > 0 \\
arbitrarily low (i.e. \leq MC^k(0)) & \text{otherwise}
\end{cases}$$

Observe that incorporating the manufacturing leaf’s optimal decision into the parent node’s optimization in this manner is equivalent to finding the sub-game perfect equilibrium for problem 1.
5.1.2 Procedure B: Retailer Elimination from Branch Point Downstream of the Leader

We will now consider elimination of multiple downstream leaves at the same time. This is necessary when one node supplies to multiple downstream leaves, i.e. at a branch point, since the quantities supplied to the downstream nodes are related to each other. Since we apply this recursively to the supply network, a “branch point” may have only one downstream child retailer. The important characteristics are that there are one or more retailer nodes, and that they all have one common upstream parent. Thus, in this procedure we will eliminate all the child nodes when the branch point is the local leader, and offers a contract to all the downstream nodes.

Let $j$ be a branch point node, i.e., $j$ is the upstream parent of $L$ leaves $\{k_l\}_{l=1}^L$. Thus $j$ both offers a contract to, and sells goods to the leaves $k_l$. Note that $j$ is free to offer different wholesale prices and sell different quantities to each child. In addition, $j$ may also offer contracts to other upstream nodes and buy goods from them (we will not consider these nodes here). This is shown in Figure 7(b). In this procedure we will eliminate the leaves $k_1..k_L$ and incorporate their marginal functions into node $j$’s optimization.

Their optimization problems are given by Problem 3.

**Problem 3** The optimization for the retailer leaves’ parent node $j$ (which itself has an upstream parent) can be rewritten as

$$\max \pi^j \left( \left\{ q^{j,k} \right\}_{k \in C_d(j)}, \ldots, \left\{ w^k \right\}_{k \in C_d(j)}, \ldots \right)$$

$$= -w^j q^j I_u(j) - \gamma(.) + \sum_{k \in C_d(j)} w^k q^k - \int_{x=0}^{q^j} MC^j(x) \, dx$$

s.t. $q^{j,k} \geq q^k \quad \forall k \in C_d(j)$,

$\Gamma$

$q^j \geq \sum_{k \in C_d(j)} q^{j,k}$.

The term $\gamma(.)$ in (15) includes $j$’s costs from contracting with other (upstream) children. The set $\Gamma$ contains the other constraints from $j$’s optimization from contracting with upstream children. Similarly, $j$’s objective function may have other parameters and decision variables that are not directly related to contracting with $k$, for example $w^j$.

The optimization for a retailer leaf $k$ (which has no children, and has an upstream parent, node $j$) can be rewritten as

$$\max \pi^k \left( q^k, w^k \right) = \int_{x=0}^{q^k} MR^k(x) \, dx - w^k q^k$$

$$- \int_{x=0}^{q^k} MC^k(x) \, dx$$

22
Lemma 7 The optimal solution to the optimization for a retailer leaf \( k \) as shown in equation (16) is

\[
q^k = \begin{cases} 
q^k : MR^k (q^k) - MC^k (q^k) = w^k & \text{if } w^k < MR^k (0) - MC^k (0) \\
0 & \text{otherwise}
\end{cases}
\]

Define a modified problem for \( j \) as shown in Problem 4. Again, note that instead of offering a contract to each \( k \), \( j \)'s problem now includes each \( k \)'s revenues and costs transformed by \( F \).

Problem 4 Rewrite the objective function for \( j \) in Problem 3 by replacing \( w^k q^k \) with \( k \)'s revenue and cost functions transformed by \( F \), i.e.

\[
\max_{\{q^k\}_{k \in C_d(j)}} \pi^j \left( \{q^k\}_{k \in C_d(j)} \right) = -w^j q^j I_u (j) - \gamma \left( . \right) + \sum_{k \in C_d(j)} \left( \int_{x=0}^{q^j} F_x \left( MR^k (x) - MC^k (x) \right) dx \right) - \int_{x=0}^{q^j} MC^j (x) dx \\
\text{s.t. } \Gamma. \quad (17)
\]

Lemma 8 The optimal solution to \( j \)'s optimization in Problem 3 (incorporating optimal behavior by \( k \)) and the optimal solution to \( j \)'s optimization in Problem 4 have identical \( q^j \), \( \{q^k\}_{k \in C_d(j)} \) and \( \pi^j \).

Corollary 3 In Problem 3, for each \( k \in C_d (j) \), \( j \) will always set \( q^{j,k} = q^k \), or, equivalently, set each \( w^k \) such that \( q^k \) equals \( q^{j,k} \).

Corollary 4 For an equilibrium solution where \( k \) produces \( q^k \) (obtained after solving the complete problem), we can calculate the wholesale price offered by \( j \) to \( k \) in problem 3 as

\[
w^k = \begin{cases} 
MR^k (q^k) - MC^k (q^k) & \text{if } q^k > 0 \\
\text{arbitrarily high (i.e., } \geq (MR^k (0) - MC^k (0)) \text{) otherwise.}
\end{cases}
\]

Again, incorporating the retailer leaves' optimal decision into the parent node's optimization in this manner is equivalent to finding the sub-game perfect equilibrium for problem 3.

5.1.3 Network Reduction Algorithm

We will now use the two node elimination procedures in order to outline the algorithm which will reduce an entire supply network into an equivalent network consisting of one node, the leader. This
happens when the leader is upstream of the first branch point in the network. The reduced form of the network is equivalent in that the same quantities will be sold in each of the markets as in the original network.

Algorithm 1 Network Reduction Algorithm:

1. If the tree has a manufacturing leaf (which is not the leader), then let that leaf be node $k$ and its parent be node $j$
   
a. use Lemma 6 from Section 5.1.1 to eliminate node $k$, and incorporate node $k$’s marginal cost function into node $j$’s optimization problem,
   
b. redefine the manufacturing leaf set $M$, and the network definitions $T$, $S$ and $C$ to reflect the node elimination, and
   
c. go back to Step 1.
   Otherwise, go to Step 2.

2. If the tree has one or more retailer leaves with a common parent, with all the retailer leaves downstream of the leader (i.e. a branch point downstream of the leader), then let the common parent be $j$ and the downstream retailer leaves be $\{k\}_{(j,k) \in C, (j,k) \in S}$
   
a. use Lemma 8 from Section 5.1.2 to eliminate the nodes $\{k\}_{(j,k) \in C, (j,k) \in S}$, and incorporate their marginal demand and cost functions into node $j$’s optimization problem
   
b. redefine the retailer leaf set $R$, and the network definitions $T$, $S$ and $C$ to reflect the node eliminations, and
   
c. go back to Step 1.
   Otherwise, Stop.

<End of Algorithm>

Lemma 9 Algorithm 1 terminates in a finite number of steps.

Each application of one of the procedures A and B incorporates the cost and revenue functions of one or more child nodes into theirs parent’s optimization. Thus, the algorithm reduces the network in a manner that the reduced network also has the same solution as the original network.

5.1.4 Sub-Game Perfect Equilibrium

For any node $j$ in the extensive form game, let $\phi(j)$ be the sub-game defined at node $j$. $\phi(j)$ contains all nodes and branches that follow $j$, including $j$ itself. Note that $j$ is a sub-root, i.e. any node acting after $j$ knows that $j$ has already decided. A subgame-perfect equilibrium of the complete extensive form game is any equilibrium in strategies such that for every subgame of the complete game, the restriction of these strategies to this subgame is also an equilibrium in strategies of the subgame (see Myerson (1991)).

Algorithm 1 successively incorporates the children’s optimal decisions into their parent’s optimization. Thus, at each stage of this algorithm, the remaining nodes’ optimizations preserves the sub-game perfect solution for the nodes eliminated thus far.
An important characteristic of the equilibrium for the subgame with root at node $j$ is that the only decision of any parent that is relevant to $j$’s optimal decision is the wholesale price $w^j$ that $j$’s parent (node $r$) offers $j$. However, the quantity that each parent selects is also important, although it does not enter the optimal decision by $j$. To illustrate this, consider $j$, and $j$’s parent node $r$. Node $r$ offers $w^j$ to node $j$. Since $r$ can be either upstream or downstream of $j$, we will consider the two cases separately.

If $r$ is upstream of $j$, then $r$’s quantity choice is constrained to $q^{r;j} \geq q^j$ (see the objective functions in section 3.4). Thus we do not allow $r$ to select any quantity that violates this constraint. Since this is a game of complete information, $r$ can accurately anticipate the consequences of any $w^j$ that she offers $j$, and can thus satisfy this constraint. Thus $r$ will always supply the quantity that $j$ orders from her.

If $r$ is downstream of $j$, then $r$’s quantity choice is constrained to $q^r \leq q^{j;r}$. Thus, while $r$ can offer any $w^j$ to $j$, she can only sell a maximum of the quantity that $j$ supplies her. It is thus feasible for node $r$ to arbitrarily select $q^r$ and $w^j$ for the subgame $\phi(j)$ in such a manner that $j$ selects $q^{j;r} > q^r$. However, once we consider the subgame $\phi(r)$, this decision is not the optimal one for $r$, and $r$ will select the optimal combination of $w^j$ and $q^r$. According to Corollaries 1, 3 and 5, this will satisfy $q^r = q^{j;r}$.

### 5.2 Existence and uniqueness of solutions

In the previous section we described the algorithm by which we reduce a network to a problem for one node. In this section we establish the existence and uniqueness of the solution for this class of networks.

We will first apply the procedure to obtain the solution to a basic network. The basic network with no upstream branch point class is shown in Figure 8 (a).

![Network Examples](Figure 8: Network Examples.)
Consider the network in Figure 8 (a). The leader is node 1. The retailers are nodes 2 and 3.

Stage 1 (Node 1):

\[
\pi^1 (w^2, w^3) = w^2 q^{1,2} + w^3 q^{1,3} - \int_0^{q^1} MC^1 (x) \, dx
\]

s.t. \( q^{1,2} \geq q^2 \)
\( q^{1,3} \geq q^3 \)
\( q^1 \geq q^{1,2} + q^{1,3} \)

Stage 2 (Nodes 2 and 3):

\[
\pi^2 (q^2; w^2) = -w^2 q^2 + \int_0^{q^2} (MR^2 (x) - MC^2 (x)) \, dx.
\]

\[
\pi^3 (q^3; w^3) = -w^3 q^3 + \int_0^{q^3} (MR^3 (x) - MC^3 (x)) \, dx.
\]

An application of the algorithm works this way. One application of Procedure B eliminates nodes 2 and 3, thus reducing the problem to the leader. Incorporating the binding constraints \( q^{1,2} = q^2 \), \( q^{1,3} = q^3 \) and \( q^1 = q^{1,2} + q^{1,3} \), the final reduced objective function for the leader, node 1 will be

\[
\pi^1 (q^2, q^3) = \begin{cases} 
\int_0^{q^2} F_{q^2} (MR^2 (x)) \, dx + \int_0^{q^3} F_{q^3} (MR^3 (x)) \, dx \\
- \left\{ \int_0^{q^2} F_{q^2} (MC^2 (x)) \, dx + \int_0^{q^3} F_{q^3} (MC^3 (x)) \, dx + \int_0^{q^2+q^3} MC^1 (x) \, dx \right\}
\end{cases}
\]

When we take the first order conditions, only those marginal cost expressions remain which apply to that particular market. There are two first-order-conditions, for \( q^2 \) and \( q^3 \). Define the following expressions (which are the derivatives of the profit function with respect to the market quantities)

\[
\Delta_{q^2} (q^2, q^3) = F_{q^2} (MR^2 (q^2)) - F_{q^2} (MC^2 (q^2)) - MC^1 (q^2 + q^3)
\]
\[
\Delta_{q^3} (q^2, q^3) = F_{q^3} (MR^3 (q^3)) - F_{q^3} (MC^3 (q^3)) - MC^1 (q^2 + q^3).
\]

The first-order-condition with respect to \( q^2 \) and \( q^3 \) (with complementary slackness) is

\[
\Delta_{q^2} (q^2, q^3) q^2 = 0
\]
\[
\Delta_{q^3} (q^2, q^3) q^3 = 0.
\]

Hence we have two equations in two unknowns. Note that in each characteristic equation the marginal function(s) for each node is(are) marginalized exactly by the number of contracts that separate it from the leader. The solution to this is composed of four regions, depending upon whether
any amount is sold in each market. The solution is

$$\left( q^2, q^3 \right) = \begin{cases} 
(0, 0) & \text{if } \Delta_{q^2} (0,0) < 0 \text{ and } \Delta_{q^3} (0,0) < 0 \\
(q^{2*}, 0) & \text{if } \Delta_{q^2} (0,0) > 0 \text{ and } \Delta_{q^3} (q^{2*},0) < 0 \\
(0, q^{3*}) & \text{if } \Delta_{q^2} (0,q^{3*}) < 0 \text{ and } \Delta_{q^3} (0,0) > 0 \\
(q^{2#}, q^{3#}) & \text{if } \Delta_{q^2} (0,q^{3*}) > 0 \text{ and } \Delta_{q^3} (q^{2*},0) > 0
\end{cases}$$

where $q^{2*}$ is the solution to $\Delta_{q^2} (q^2,0) = 0$

and $q^{3*}$ is the solution to $\Delta_{q^3} (0,q^{3}) = 0$

and $\left( q^{2#}, q^{3#} \right)$ is the simultaneous solution to the two equations in (21).

**First-Order Conditions for the Network without an Upstream Branch Point**  For a network which uses only procedures A and B, we can reduce the problem to one optimization for the leader.

Thus, the leader’s optimization will have the following first-order-conditions (one for each market)

$$F_{q^r}^{\delta(r;j)} \left( MR^r \left( q^r \right) \right) - \sum_{j \in S^r} F_{q^r}^{\delta(i,j)} \left( MC^j \left( \sum_{k \in R^j} q^k \right) \right) = 0 \quad \forall r \in R. \quad (22)$$

**Lemma 10**  There exists a unique solution to the reduced problem for networks where the leader is upstream of the first branch point.

Thus, the solution to the problem may lie in a number of distinct regions, each of which is defined by which $q^r$ is non-zero. For a problem with $|R|$ markets, there are $2^{|R|}$ solution regions.

6 Networks with Upstream Branch Points

We will now consider the networks in which there is a branch point upstream of the leader. This is more involved than the networks presented thus far. We will first present some examples of networks with upstream branch points.

6.1 Upstream Branch Point Examples

Consider two examples as shown in Figure 9. The first is a network where we have two nodes between the branch point and the leader, shown in Figure 9 (a). The second example is similar. We introduce another node between the leader and the first branch point, but this new node additionally serves another retailer in addition to the leader, shown in Figure 9 (b). We cannot reduce these network any further using only Procedures A and B defined earlier.

We will define a third procedure, Procedure C, to solve the class of networks shown in Figure 9 (c). This is a network in which the leader is immediately downstream of the first branch point. The first branch point may have other upstream nodes, and the nodes downstream of the first branch point may
have other downstream nodes and branch points. Similarly, the leader may have other downstream nodes and branch points, as well as other upstream nodes. Applications of Procedure A reduces the network upstream of the first branch point and upstream of the leader, and applications of Procedure B reduce the network downstream of the leader as well as downstream of the nodes downstream of the first branch point. Subsequently, Procedure C will allows us to eventually reduce this network to a problem over one node. Once we have done this, we indicate how the network reduction algorithm can be extended to solve the examples in Figures 9 (a) and (b) as well.

To illustrate these methods we will first show an example using a 3-node network where there is a branch point upstream of the leader.

6.2 Basic Network with an Upstream Branch Point

6.2.1 Network Reduction

We will now consider the solution for a basic network with one upstream branch point (the smallest such network consists of three nodes). This basic network is shown in Figure 8 (b). The common supplier is node 1. The leader is node 2. The retailers are nodes 2 and 3. The three-stage game is as follows:

\[
\begin{align*}
\pi^2(q^2, w^1) &= -w^1 q^{1,2} + \int_0^{q^2} (MR^2(x) - MC^2(x)) \, dx \\
\text{s.t.} & \quad q^2 \leq q^{1,2}
\end{align*}
\]

Figure 9: Examples of upstream branch point networks.
Stage 2 (Node 1):

\[ \pi^1(q^{1,2}, w^3; w^1) = w^1 q^{1,2} + w^3 q^{1,3} - \int_{0}^{q^1} MC^1(x) \, dx \]

s.t. \[ q^{1,3} \geq q^3 \]
\[ q^1 \geq q^{1,2} + q^{1,3} \]

Stage 3 (Node 3):

\[ \pi^3(q^3, w^3) = -w^3 q^3 + \int_{0}^{q^3} (MR^3(x) - MC^3(x)) \, dx. \]

Recollect that \( q^{1,2} \) is the quantity that 1 supplies 2. We can eliminate node 3 using a reduction method similar to Procedure B (this is formally defined as procedure C later), reducing the problem to a two-stage problem between the leader and the branch point. Incorporating the binding constraints \( q^{1,3} = q^3 \) and \( q^1 = q^{1,2} + q^{1,3} \), we obtain:

Stage 1 (Node 2):

\[ \pi^2(q^2, w^1) = -w^1 q^{1,2} + \int_{0}^{q^2} (MR^2(x) - MC^2(x)) \, dx \]

s.t. \( q^2 \leq q^{1,2} \)

Stage 2 (Node 1):

\[ \pi^1(q^{1,2}, q^3; w^1) = w^1 q^{1,2} + \int_{0}^{q^3} F_x (MR^3(x) - MC^3(x)) \, dx - \int_{0}^{q^{1,2}+q^3} MC^1(x) \, dx. \]

The second stage problem solution involves determining the relationship between the wholesale price that retailer 2 offers the common supplier 1, and the quantity that the supplier sells to retailer 2, i.e., \( q^{1,2} (w^1) \).

**Stage 2 Solution (Node 1)** We will first solve node 1’s problem. The first-order-condition for this is

\[ \frac{\partial \pi^1}{\partial q^{1,2}} = w^1 - MC^1(q^{1,2} + q^3) = 0 \]
\[ \frac{\partial \pi^1}{\partial q^3} = F_q^3 (MR^3(q^3)) - \{ F_q^3 (MC^3(q^3)) + MC^1(q^{1,2} + q^3) \} = 0. \]

Observe that the solution can be written in terms of the two first order conditions (along with complementary slackness):
\[(w^1 - MC^1(q^{1,2} + q^3)) q^{1,2} = 0\]
\[(F_{q^3} (MR^3(q^3)) - \{F_{q^3} (MC^3(q^3)) + MC^1(q^{1,2} + q^3)\}) q^3 = 0.\]

Thus, node 1’s optimal decision (in the second stage) is

\[q^{1,2}, q^3 = \begin{cases} 
(q^{1,2*}, 0) & \text{if } w^1 > F_{q^3} \{MR^3(q^3) - MC^3(q^3)\} \big|_{q^3=0} \\
(0, q^{3*}) & \text{if } w^1 < F_{q^3} \{MR^3(q^3) - MC^3(q^3)\} \big|_{q^3=q^3*} \\
(q^{1,2#}, q^{3#}) & \text{otherwise}
\end{cases}\]

where

\[q^{1,2*} \text{ solves } w^1 - MC^1(q^{1,2}) = 0\]

\[q^{3*} \text{ solves } (F_{q^3} (MR^3(q^3)) - \{F_{q^3} (MC^3(q^3)) + MC^1(q^{1,2} + q^3)\}) q^3 = 0\]

\[(q^{1,2#}, q^{3#}) \text{ solves } \begin{cases} 
  w^1 - MC^1(q^{1,2} + q^3) = 0 \\
  w^1 - (F_{q^3} (MR^3(q^3) - MC^3(q^3))) = 0 
\end{cases}.\]

Thus, the optimal solution to node 1’s problem comprises of $q^{1,2}(w^1)$ and $q^3(w^1)$ as a function of the parameter $w^1$. If the wholesale price $w^1$ is high enough, that is, if it is higher than the marginalized contribution from selling the first unit to retailer 3, then node 1 will not sell anything to node 3, and instead concentrate on selling everything to retailer 2. In this case, the network effectively collapses to a two-node network where retailer 3 is ignored.

On the other hand, if the wholesale price $w^1$ is lower than the optimal marginalized contribution from selling the optimal quantity to retailer 3 alone, then node 1 concentrates on selling only to retailer 3. In this case we have the unusual situation that the network collapses to the two-node network without the original leader, and node 1 as the effective leader. The next Remark highlights this observation.

**Remark 1** If $w^1$ is lower than the contribution (to the common supplier) from selling the optimal quantity to retailer 3 alone, then the common supplier does not sell to the leader, node 1, i.e.

\[q^{1,2} = 0 \text{ if } w^1 < F_{q^3} \{MR^3(q^3) - MC^3(q^3)\} \big|_{q^3=q^{3*}}\]

where $q^{3*}$ solves $(F_{q^3} (MR^3(q^3)) - \{F_{q^3} (MC^3(q^3)) + MC^1(q^3)\}) q^3 = 0$.

Finally, if the wholesale price is neither too high nor too low, then node 1 will sell to both retailers. In this case, we expect the quantity sold to each retailer is less than what it would be if node 1 was selling exclusively to that retailer.

**Stage 2 Solution with linear MC and MR** In order to solve the problem for the next stage, it is more convenient to consider the model with linear marginal functions. We define the marginal cost for node $k$ as

\[MC^k(q) = 2K^k q + c^k.\]
when it manufactures/processes a quantity \( q > 0 \), where \( K^k, c^k \geq 0 \). Recollect that we require an increasing marginal cost for all manufacturers, i.e., \( K^k \) has to be positive for all manufacturing nodes, thus
\[
K^k > 0 \quad \forall k : k \in M.
\] (24)
Similarly, the linear demand function is
\[
p^r(q) = a^r - b^r q
\] (25)
where \( a^r, b^r > 0 \). Let
\[
MR^r(q) = \frac{\partial}{\partial q} q p^r(q) = a^r - 2b^r q.
\] (26)

With linear marginal functions, node 1’s optimal decision (in the second stage) is
\[
(q^{1.2}, q^3) = \begin{cases} 
(q^{1.2*}, 0) & \text{if } w^1 > (a^3 - c^3) \\
(0, q^{3*}) & \text{if } w^1 < (a^3 - 4b^3q^{3*}) - (c^3 + 4K^3q^{3*}) \\
(q^{1.2#}, q^{3#}) & \text{otherwise}
\end{cases}
\]
where
\[
q^{1.2*} = \frac{w^1 - c^3}{2K^3}
\]
\[
q^{3*} = \text{max} \left( 0, \frac{a^3 - c^3 - c^1}{4b^3 + 4K^3 + 4K^3} \right)
\]
\[
(q^{1.2#}, q^{3#}) \text{ solves } \begin{cases} 
q^1 - (c^1 + 2K^1 (q^{1.2} + q^3)) = 0 \\
q^1 - ((a^3 - 4b^3q^3) - (c^3 + 4K^3q^3)) = 0
\end{cases}
\]

Note that the “kink” in \( q^{1.2} (w^1) \) occurs when \( w^1 = a^3 - c^3 \), i.e., when retailer 2 offers a wholesale price that is exactly equal to the profitability of the first unit that can be sold in retailer 3’s market.

**Stage 1 Solution (Node 2)**

Now, the leader, node 2’s objective function (in the first stage) is
\[
\pi^2 (q^2, w^1) = -w^1 q^{1.2} + \int_0^{q^2} \left\{ MR^2(x) - MC^2(x) \right\} dx \quad \text{s.t. } q^2 \leq q^{1.2}
\]
\[
= \begin{cases} 
0 & \text{if } q^{1.2} = 0 \\
-MC^1(q^{1.2} + q^3) q^{1.2} + \int_0^{q^2} \left\{ MR^2(x) - MC^2(x) \right\} dx & \text{if } q^{1.2} > 0
\end{cases}
\]

Note that \( q^{1.2} = q^{1.2} (w^1) \) and \( q^3 = q^3 (w^1) \).

Note that if \( w^1 \) is low, then \( q^{1.2} \) is zero, and \( \pi^2 \) is zero too. Otherwise, for a given \( w^1 \), we first obtain \( q^{1.2} \) and \( q^3 \) from node 1’s optimal solution, and then solve for \( q^2 \). Recollect that \( q^3 \) is zero if \( w^1 \) is very high.
6.2.2 The Cost Function for the Leader

We can now examine the effective cost function for the leader, node 2. The following Lemmas build upon the analysis for the basic network.

**Lemma 11** \( q^2 = q^{1.2} \).

We can incorporate \( q^2 = q^{1.2} \) into the optimization, and obtain

\[
\pi^2(w^1) = -MC^1(q^2 + q^3)q^2 + \int_0^1 \{MR^2(x) - MC^2(x)\} \, dx
\]

where \( q^2 = q^{1.2}(w^1) \).

**Lemma 12** \( w^1(q^2) \) is increasing and convex. With linear marginal functions we have

\[
w^1(q^2) = \begin{cases} 
  w^1_D + \frac{K^1(4K^3 + 4K^3)}{K^1 + 2b^3 + 2K^3}q^2 & \text{when } q^3 > 0 \\
  c^1 + 2K^1q^2 & \text{when } q^3 = 0
\end{cases}
\]

where \( w^1_D \) is given by \( w^1_D = F(MR^3(q^3) - MC^3(q^3)) \). With linear marginal functions this is

\[
w^1_D = (a^3 - c^3) \left( \frac{K^1}{b^3 + K^3 + K^1} \right) + \frac{b^3 + K^3}{b^3 + K^3 + K^1}c^1.
\]

**Lemma 13** \( F(q^2 \{ w^1(q^2) \}) \) has the following properties

(a) it is increasing with a break point corresponding to the kink point of \( w^1(q^2) \),

(b) the slope increases at the break point, i.e., at any two points \( x_A, x_B \in [0, \infty) \), where \( x_B > x_A \), the slope at \( x_B \) is not less than the slope at \( x_A \) (if both exist).

We illustrate this in Figure 10. This figure shows the relationship between \( w^1 \) and \( q^2 \), assuming that the retailer 3’s market is profitable enough to be served in absence of retailer 2 (i.e. \( a^3 > c^1 + c^3 \)).

From our analysis, when node 1 sells to both markets \( \frac{dq^2}{dw^1} \) is higher, i.e. \( \frac{dq^2}{dw^1} \bigg|_{q^3=0} < \frac{dq^2}{dw^1} \bigg|_{q^3>0} \). One way to explain this is as follows: when the common supplier is selling to both markets, if the leader increases \( w^1 \) by a small amount, this not only increases the quantity that the common supplier would have in any case supplied the leader, but it also diverts more supplies from the other market, since the leader is now offering a more profitable wholesale price. This means that \( w^1 \) as a function of \( q^2 \) is continuous and piecewise linear, and as earlier, it is still increasing and convex (although it is no longer differentiable everywhere). Thus, in the graph of \( w^1 \) as a function of \( q^2 \), the slope is smaller when node 1 sells to both retailers (compared to when node 1 sells to node 2 exclusively). The line ABC is \( MC^1(q^2) \), which is the relationship between \( w^1 \) and \( q^2 \) in the presence of retailer 3. The relationship of \( w^1(q^2) \) in the presence of retailer 3 is given by the line DBC. Thus, the presence of the other retailer increases the procurement cost for retailer 2. Point A is obtained from \( w^1_A = MC^1(0) \).

Point D is determined by \( w^1_D = F(MR^3(q^3) - MC^3(q^3)) \), i.e. the minimum \( w^1 \) at which supplier 1 will sell to retailer 2, which occurs when \( w^1 \) equals the (transformed) optimal marginal
Figure 10: The relationship between $w^1$ and $q^2$, and the implied marginal cost function.

revenue from retailer 3 in the absence of retailer 2 (note that $q^{3\ast}$ is the optimal solution when supplier 1 sells only to retailer 3). Finally, the value of $q^2$ for point B can be calculated from $F \left( MR^3 (\epsilon) - MC^3 (\epsilon) \right) = w^1_B = MC^1 (q^2 + \epsilon)$ where $\epsilon$ is very small. In other words, as $q^2$ increases, the last unit sold to retailer 3, which has the highest profitability (since it is the “first” unit sold in retailer 3’s market) must have a transformed marginal revenue equal to the marginal cost at the volume $q^2$. An interesting observation is that as $MC^1 (\cdot)$ becomes “flatter”, the value of $q^2$ at which the kink occurs becomes very high, and the optimal solution occurs to the region where the supplier 1 sells to both retailers.

The implied marginal cost function for retailer 2 is given by $F_{q^2} \left( w^1 (q^2) \right)$, which is DEFG in Figure 10, with the line segment starting from F is open ended (i.e. DEFG is lower semi-continuous). Note that when supplier 1 sells to more than one other retailer, e.g. retailers 3, 4 and more, then the line DBC changes so that $w^1 (q^2)$ is now piecewise linear, increasing and convex. For small values of $w^1$, supplier 1 does not sell to retailer 2. As $w^1$ increases, supplier 1 finds it profitable to sell to retailer 2, and as $w^1$ increases higher, supplier 1 gradually stops selling to some other retailers and sells more to retailer 2. A kink exists in the function whenever supplier 1 stops selling to a retailer. The transformed function has a discontinuous jump at each of these kinks. The number of kinks equals the number of non-identical retailers that supplier 1 sells to in the absence of retailer 2. (Retailer $x$ is “identical” to retailer $y$ if $MR^x (q) - MC^x (q) = MR^y (q) - MC^y (q)$.)

In the next section we define Procedure C, which allows us to reduce the network at a branch point upstream of the leader. This results in the leader obtaining a piecewise linear marginal cost function.
It is not possible to obtain the cost function in closed form.

The next step in the analysis of the supply network would be an application of Procedure A to eliminate node 1, and thus leave a reduced optimization for the leader node alone. This means that the node 2’s problem has the marginalized function $F_{q^2}(w^1(q^2))$ (in order to say this, we also need to extend the definition of $F_x(f(x))$, which we after defining Procedure C).

**Lemma 14** $\pi^2$ has a unique maximum after incorporating the optimal decision by node 1, i.e., $q^2(w^1)$, given by

$$(MR^2(q^2) - MC^2(q^2) - F_{q^2}(w^1(q^2)))q^2 = 0.$$ 

The solution $q^2$ is positive if

$$MR^2(0) - MC^2(0) > F_{q^2}(w^1(q^2))|_{q^2=0}$$

which reduces to

$$a^2 - c^2 > w_D^1 \text{ (from Lemma 12)}$$

where $w_D^1 = (a^3 - c^3)\left(\frac{K^1}{b^3 + K^3 + K^1}\right) + \frac{b^3 + K^3 + K^1}{b^3 + K^3 + K^1}c^1$.

From the Figure 10 we can observe that the solution depends upon the intersection of the marginal revenue function with $F(q^2)$. On this graph, $MR^2(q^2) = a^2 - b^2q^2$ will be a decreasing linear function. Thus, $MR^2(q^2)$ may intersect DEFG in three possible regions. If the intersection lies in region FG, this means that market 2 is so profitable that market 3 is not served at all. If the intersection lies in region EF, this means that market 2 is still profitable enough to shut out market 3, but just barely. (This also explains the reason why $F_x(f(x))$ is later defined to be lower semi-continuous for increasing $f$.) Finally, if the intersection lies in region DE, then both markets are served. It is also possible that $MR^2(0)$ lies below the point D on the vertical axis. This means that market 2 will not be served at all. In that case, if the point D is above the point A, market 3 will be served. If not, then neither market is served.

**6.3 Procedure C**

We will now formally define Procedure C. Observe that we will only be able to reduce the network by eliminating the downstream child nodes, but not the branch point itself. The previous example using a 3-node network illustrates the principle; it is not possible to obtain closed form expressions for the marginalized cost and revenue functions across a branch point upstream of the leader.
6.3.1 Procedure C: Retailer Elimination from Branch Point Upstream of the Leader

We consider elimination of child nodes at a branch point when one of the downstream nodes is the local leader. This is depicted in Figure 7 (c).

Let $j$ be a branch point node, i.e., $j$ is upstream of $L$ nodes $\{k_l\}_{l=1}^L$. $k_1$ is the parent of $j$, and $j$ is the parent of all other $k_l$ nodes $(l = 2..L)$. Thus, $j$ sells to all downstream nodes, but receives a contract from $k_1$ and offers contracts to all other $k_l$. Note that $k_1$ may also offer contracts to other nodes and buy/sell goods from them (we will not consider these nodes here). In this procedure we will eliminate nodes $k_2..k_L$, and incorporate their marginal functions into $j$’s problem.

Observe that $j$ simultaneously supplies all the downstream nodes whether receiving a contract from such a node or offering a contract to such a node. However, since $j$ offers a contract to the nodes $k_2..k_L$ after receiving a contract from $k_1$, it follows that if $j$ supplies to more than one downstream node, then all of these must be equally profitable. The contract received from $k_1$ (i.e., $w_j$) cannot be changed by $j$, but the contracts offered to $k_2..k_L$ can be changed by $j$. Thus, it is possible that supplying one or more of $k_2..k_L$ is more profitable than supplying to $k_1$. In such a situation, $j$ will not supply $k_1$. This will happen if $w_j$ is low enough. It is also possible that some of the contracts with $k_2..k_L$ can be profitable than the others (e.g. because some of the markets are larger). In that case, $j$ may choose to supply only some of its downstream children and not others.

Note that we will assume that $j$ has supply and contract relationships only to $k_1..k_L$. In a given network $j$ may also procure from other manufacturers. However, our algorithm to reduce the network ensures that we will consider a leader-upstream-branch-point elimination only after having eliminated these other nodes. Hence the procedure does not need to consider more complicated structures.

Their optimization problems are given by Problem 5.

Problem 5 Stage 1: Node $k_1$’s optimization is

$$\begin{align*}
\max & \quad \pi^{k_1} (..., w^j, ...) = \gamma(.) - w^j q^{j,k_1} \\
\text{s.t.} & \quad q^{k_1} \leq q^{j,k_1} \\
& \quad \Gamma.
\end{align*} \quad (27)$$

The term $\gamma(.)$ in equation (27) includes $k_1$’s costs and revenues from contracting with children other than $j$. The set $\Gamma$ contains the other constraints from $k_1$’s optimization from contracting with children other than $j$.

Stage 2: The optimization for the branch node $j$ (which has a downstream parent $k_1$) can be rewrit-
\[
\max \pi^j \left( \left\{ q^{j,k}, w^k \right\}_{k \in C_d(j)}, q^{j,k_1}; w^j \right) = w^j q^{j,k_1} + \sum_{k \in C_d(j)} w^k q^k - \int_{x=0}^{q^j} MC^j(x) \, dx
\]

\text{s.t.} \quad q^{j,k} \geq q^k \quad \forall k \in C_d(j),
q^j = q^{j,k_1} + \sum_{k \in C_d(j)} q^k.

Since the procedure by which we eliminate nodes first eliminates all upstream children, node j’s optimization is completely specified here.

Stage 3: The optimization for a retailer leaf \( k \in C_d(j) \) (each \( k \) has no children, and has an upstream parent, node j) can be rewritten as

\[
\max \pi^k \left( q^k; w^k \right) = \int_{x=0}^{q^k} MR^k(x) \, dx - w^k q^k - \int_{x=0}^{q^k} MC^k(x) \, dx
\]

Lemma 15 The optimal solution to the optimization for a retailer leaf \( k \) as shown in equation (29) is

\[
q^k = \begin{cases} 
\{ q^k : MR^k(q^k) - MC^k(q^k) = w^k \} & \text{if } w^k < MR^k(0) - MC^k(0) \\
0 & \text{otherwise}
\end{cases}
\]

Define a modified problem for \( j \) and \( k_1 \) as shown in Problem 6.

\textbf{Problem 6} We have \( k_1 \)’s problem (unchanged) in Stage 1:

\[
\max_{\ldots,w_1,\ldots} \pi^{k_1} (\ldots,w^j,\ldots) = \gamma(\ldots) - w^j q^{j,k_1}
\]

\text{s.t.} \quad q^{k_1} \leq q^{j,k_1}
\Gamma.

We have \( j \)’s problem in Stage 2:

\[
\max_{q^{j,k_1},\{q^k\}_{k \in C_d(j)}} \pi^j \left( \left\{ q^{j,k_1}, \left\{ q^k \right\}_{k \in C_d(j)} \right\}; w^j \right) = w^j q^{j,k_1} + \sum_{k \in C_d(j)} \left( \int_{x=0}^{q^k} \left\{ MR^k(x) - MC^k(x) \right\} \, dx \right) - \int_{x=0}^{q^j} MC^j(x) \, dx
\]

\text{s.t.} \quad q^j = q^{j,k_1} + \sum_{k \in C_d(j)} q^k
\Gamma.
Lemma 16 The optimal solution to j’s optimization in Problem 5 (incorporating optimal behavior by k) and the optimal solution to j’s optimization in Problem 6 have identical $q^j, \{q^k\}_{k \in C_d(j)}$ and $\pi^j$.

Observe that in this case we cannot completely reduce the problem, but instead define it as a two stage problem, where in the second stage node $j$ takes $w^j$ as an exogenous state variable, and in the first stage node $k_1$ uses the relationship between $w^j$ and $q^{jk_1}$ obtained from stage 2. (Recollect that further analysis for the basic network was illustrated in Section 6.2.)

Corollary 5 In Problem 5, for each $k \in C_d(j)$, $j$ will always set $q^{jk} = q^k$, or, equivalently, set $w^k$ such that $q^k$ equals $q^{jk}$.

Corollary 6 For an equilibrium solution where $k$ produces $q^k$ (obtained after solving the complete problem), we can calculate the wholesale price offered by $j$ to $k \in C_d(j)$ in problem 5 as

$$w^k = \begin{cases} MR^k (q^k) - MC^k (q^k) & \text{if } q^k > 0 \\ \text{arbitrarily high} & \text{otherwise} \end{cases}.$$

Again, incorporating the child retailer leaves’ optimal decision into the branch point node’s optimization in this manner is equivalent to finding the sub-game perfect equilibrium for stages 2 and 3 in problem 5.

Lemma 17 If the common supplier $j$ sells to more than one downstream child, for any two children $k_a$ and $k_b$ where $k_a, k_b \in C_d(j)$

$$F_{q^{ka}} \left( MR^{ka} (q^{ka}) - MC^{ka} (q^{ka}) \right) = F_{q^{kb}} \left( MR^{kb} (q^{kb}) - MC^{kb} (q^{kb}) \right).$$

If she sells to any downstream child $k \in C_d(j)$ as well as to the parent $k_1$, then

$$F_{q^k} \left( MR^k (q^k) - MC^k (q^k) \right) = w^j.$$

Lemma 17 states that the common supplier will price discriminate among the different child nodes. This can be seen from the fact that the wholesale prices given in Lemma 6 are $MR^k (q^k) - MC^k (q^k)$, but in order to determine quantities the common supplier equates $F_{q^k} \left( MR^k (q^k) - MC^k (q^k) \right)$, i.e. the marginalized marginal revenues.

Corollary 7 Let

$$w_{\text{max}} = \max_k \left( MR^k (q^k) - MC^k (q^k) \right) \bigg|_{q^k=0}.$$

For convenience, supplying $k_1$ can be represented by an induced marginal revenue function which is flat, with a value of $w_j$ independent of quantity.

(a) The common supplier $j$ will supply no retailer if

$$MC^j (0) \geq w_{\text{max}}.$$

(b) The common supplier $j$ will not supply those $k \in \{k_2, k_L\}$ for whom

$$w^j \geq MR^k (0) - MC^k (0).$$
$w_{\text{max}}$ is the highest revenue that $j$ can obtain from selling a unit, since it is the most profitable one. If $w_{\text{max}}$ is smaller than the marginal cost intercept $MC_j(0)$, then $j$ supplies nobody. After elimination of the children retailers $k_2..k_L$, $j$ obtains the induced demand function (and hence the induced marginal revenue function) for each market $k_2..k_L$, and already knows the wholesale price $w_j$ at which she can supply $k_1$. If there are any induced marginal revenue functions with intercepts below $w_j$, then these markets are never supplied.

In this section we have show how Procedure C eliminates all the child retailer nodes from the branch point. In order to reduce the network further, we need to examine the cost function at the branch point $j$ when considered by the node $k_1$. In the preceding section we have established that $w^j(q^{j,k_1})$ is an increasing function but it is not differentiable everywhere. We can use the Procedures A and B to reduce the network further as long as we generalize the transformation to apply to these situations.

6.4 Transformation Definition Revisited

For the class of networks with a branch point upstream of the leader, the induced marginal cost function at the branch point is not differentiable everywhere. Thus we must extend the transformation so that the transformed function does not remain undefined at these points, by constructing the transformed function so that it is semi-continuous. Observe that we need the extended definition of the transformation after we have applied procedure C once, and then must reduce the network further (for example, in the basic network analyzed in section 6.2, we need this extended definition when we attempt to reduce the network by eliminating the common supplier, node 1).

**Definition 1** Let $f$ be a real function mapping $\mathbb{R}$, the real line, to itself, with the properties that

(a) either $f$ is increasing and lower semi-continuous, or $f$ is decreasing and continuous,

(b) $f$ is differentiable almost everywhere,

(c) for all $x_1, x_2 \in \mathbb{R}^+$ such that $0 \leq x_1 \leq x_2$,

(i) if $f$ is increasing and the derivatives exist then $\frac{df}{dx}\bigg|_{x=x_1} \leq \frac{df}{dx}\bigg|_{x=x_2}$

(ii) if $f$ is decreasing then $f$ is concave, i.e. $\frac{df}{dx}\bigg|_{x=x_1} \geq \frac{df}{dx}\bigg|_{x=x_2}$.

Observe that if a function $f$ is continuous and $n$-times differentiable everywhere, then the property (c) means that it is either increasing and convex or it is decreasing and concave.

Define the transformation $F_x(.)$ which operates on $f$ in the following manner

$$F_x (f(x)) = \begin{cases} 
    f(x) + x \frac{\partial f(x)}{\partial x} & \text{if } f(x) \text{ is differentiable at } x \\
    f(x) + x \lim_{x \to x^-} \frac{\partial f(x)}{\partial x} & \text{if } f(x) \text{ is not differentiable at } x.
\end{cases}$$

Note that $f$ is differentiable almost everywhere, therefore the left derivative exists everywhere. Since a discontinuity in $\frac{\partial f(x)}{\partial x}$ occurs only for marginal cost functions (which are increasing), we take the left
derivative in order to make $F_x(f(x))$ lower semi-continuous. Thus, if the function $f$ is increasing, with points where $f$ is not differentiable, $F_x(f(x))$ is now lower semi-continuous. The reason for constructing $F_x(f(x))$ to be lower semi-continuous has been explained in section 6.2.

Figure 11 shows some examples of increasing functions $f$, and the transformed functions.

![Graphs](image)

**Figure 11:** Examples of increasing functions and their transformed functions.

### 6.4.1 Extending the Algorithm

We have applied Procedure C to solve networks with the structure shown in Figure 9 (c), where successive application of Procedures A and B reduce the network to a form where a single application of Procedure C leads to the solution. However, by extending Algorithm 1, we can now consider all other networks as well. Extend Algorithm 1 in the following manner (in all reduction procedures we will now use the extended definition of the transformation function).

**Algorithm 2 Network Reduction Algorithm:**

1. (as in Algorithm 1)
2. (as in Algorithm 1 up to step 2b)
   - c. go back to Step 1.
   - Otherwise go to Step 3.
3. If the tree has multiple retailer leaves with a common upstream supplier, where one of the retailer leaves is the local leader (i.e. a branch point upstream of the leader), then let the common upstream supplier be $j$, the downstream parent be $k_1$ and the other downstream retailer leaves be $\{k\}_{(j,k)\in C, (j,k)\in S}$
   - a. use Lemma 16 from Section 6.3.1 to eliminate the nodes $\{k\}_{(j,k)\in C, (j,k)\in S}$, and incorporate their marginal demand and cost functions into node $j$’s optimization problem.
b. redefine the retailer leaf set $R$, and the network definitions $T$, $S$ and $C$ to reflect the node eliminations, and
c. go back to Step 1.

For example, for the network in Figure 9 (a), after one application of Procedure C (which eliminates node 3) we are left with a 3-node network, which starts from the branch point node 1 and ends at the retailer 4. We now apply Procedure A twice to successively eliminate nodes 1 and 2. An important caveat here is that it is possible that the reduced problem for the leader now has a quantity solution of zero. This simply means that the nearest upstream branch point, node 1, does not supply anything to the branch leading to the leader. However, node 1 may still supply node 3. In order to determine that, we can examine the solution to node 1’s problem when offered a wholesale price of zero by node 2. This is equivalent to solving the truncated network obtained by dropping all the nodes from the branch point to the leader, as well as the corresponding arcs.

Finally, for the network in Figure 9 (b), after one application of Procedure C we have eliminated node 3, and obtained a piecewise linear marginal cost function for node 1. We now use Procedure A to eliminate node 1 and incorporate its marginal cost function into node 2. However, at this stage we will again have to apply Procedure C to eliminate node 5, since now node 2 is an upstream branch point. Finally, an application of Procedure A eliminates node 2.

7 The Two-Part Tariff Contract (Coordinating) Transformation

Finding the coordinating transformation is equivalent to finding a contract such that the transformation is the identity transformation $I$, i.e. it leaves the parameters of the follower’s optimization unchanged while still allowing the fold-back into the leader’s problem. Thus $I(f(x)) = f(x)$ for any real function $f(x)$ which maps $\mathbb{R}$ onto itself.

In Appendix 2 we show that using two-part tariff contracts does exactly this. Thus, if members in the network contract using two-part tariffs, we use Algorithm 2, but instead of using Lemmas 6, 8 and 16, we now use Lemmas 19, 20, and 22 respectively. Thus, each elimination leaves the follower’s marginal functions unchanged. This means that the reduced optimization problem is equivalent to that for the centralized model shown in Section 3.5, i.e. the reduced problem for the leader is

$$\max_{\{q^r\}_{r \in R}} \pi^i \left( \{q^r\}_{r \in R} \right) = \sum_{r \in R} \left( \int_0^{q^r} MR^r (x) \, dx \right) - \sum_{j=1}^n \left( \int_0^{q^j} MC^j (x) \, dx \right)$$

where $q^r \geq 0$, $q^j = \sum_{r \in R} q^r$.

Hence the profit function $\pi^i$ here is also concave, and has a unique solution (see Lemma 1).

An important aspect of the coordinating contract which we have not modelled explicitly is the presence of reservation profit levels for each member, i.e. the profit level below which that member
will refuse to participate. For analytic convenience we may assume this to be zero. However, the reservation profits may also be defined to be the profits the member makes in the decentralized chain. Since the profits from the centralized chain are always higher, there is always some excess profits which can drive the two-part tariff contracting process.

The coordinating contract solution is identical to the centralized solution whether or not there is a branch point upstream of the leader. In this respect the coordinating contract is easier to analyze than the wholesale price contract. An added advantage of using two-part tariff contracts is that we also avoid the outcomes where the leader is no longer a part of the functional supply chain. Thus, in all cases the leader is able to extract the excess profits.

Note that it is possible that the centralized and decentralized models sell in different sets of markets. Thus, when the decentralized supply chain is coordinated, it may stop selling in some markets and start selling in others. An interesting situation arises if the centralized solution is such that the leader buys/sells nothing. This may happen if the leader is downstream of a branch point, but that entire branch containing the leader sells nothing. In that case, the model stipulates that the leader will still offer a two-part tariff to the branch point, but since they buy or sell no quantity, this merely involves a side payment of the excess profits from the branch point to the leader. Theoretically, the branch point is indifferent since she still makes the reservation profits; however, a much more likely scenario is that the branch point stops making the leader a side payment, and becomes the operational leader of the supply chain.

Finally, note that if two-part tariffs are used only in some parts of the contract tree, this simply reduces the number of marginalizations $F$ and replaces some of them with $I$. This may happen if it is not possible to coordinate the entire tree.

8 Conclusion

In this paper we have proposed a framework to analyze large network supply chains using wholesale price and two-part tariff contracting. The framework allows any member of the supply chain to be the contract initiator for the entire network. The basic mechanism by which we analyze network supply chains is by using a transformation which incorporates a child node’s demand and costs into the parent nodes in a manner which leaves the quantity solution unchanged. Using this we show how to reduce the entire network into one problem for the contract leader. The presence of a common supplier upstream of the leader reduces the contracting power of the leader; in some cases the leader may not even be part of the decentralized supply chain solution, since the common supplier may choose not to sell anything to him. We also show that it is always possible to coordinate the supply chain by using two-part tariff contracts. The solution is identical to the centralized solution whether or not the nominal leader buys or sells a positive quantity.

There are many examples of potential applications of this framework to models where authors have studied (a) two-tier supply chains, or (b) only assembly-type supply chains, or (c) only distribution-
type supply chains, or (d) only a specific member as the leader. Thus, this framework may be used to extend such research so that multiple tiers, combined assembly and distribution, and different leaders in the supply chain can be studied.

References


9 Appendix 1: Proofs

9.1 Concavity of the Reduced Objective Function

Consider either the centralized problem, or the reduced problem for the leader in a network with no branch points upstream of the leader.

Suppose that the final reduced problem is

\[
\max \pi^i \left( \{q_r^r\}_{r \in R} \right) = \sum_{r \in R} \left( \int_0^{q^r} \rho^r (x) \, dx \right) - \sum_{j} \left( \int_0^{q^j} \xi^j (x) \, dx \right)
\]

s.t. \(q^r \geq 0\)

and \(q^j = \sum_{r \in R^j} q^r\) (or conversely, \(q^r = \sum_{j \in S^r} q^j\))

where \(q^r\) is the quantity sold in each market \(r \in R\), \(\rho^r (x)\) is the marginal revenue function for market \(r\) (\(\rho^r (0) > 0, \frac{d}{dx} \rho^r (x) < 0\)), and \(\xi^r (x)\) is the marginal cost function for the quantity processed at node \(j\) (\(\xi^r (0) \geq 0, \frac{d}{dx} \xi^r (x) > 0\)).

Lemma 18 The objective function shown in equation (33) is concave.

Proof. of Lemma 18: Consider a point \(A = \{q_A^r\}\) and a point \(B = \{q_B^r\}\). To establish concavity from first principles, we have to show that

\[
\pi^i \left( \{\alpha q_A^r + (1 - \alpha) q_B^r\}_{r \in R} \right) \geq \alpha \pi^C \left( \{q_A^r\}_{r \in R} \right) + (1 - \alpha) \pi^C \left( \{q_B^r\}_{r \in R} \right)
\]

where \(\alpha \in [0, 1]\)

Note that for \(X \in \{A, B\}\) we have

\[
q_X^j = \sum_{r \in R^j} q_X^r
\]

and therefore \(\alpha q_A^j + (1 - \alpha) q_B^j = \alpha \sum_{r \in R^j} q_A^r + (1 - \alpha) \sum_{r \in R^j} q_B^r\)

Hence

\[
\pi^i \left( \{\alpha q_A^j + (1 - \alpha) q_B^j\} \right) = \sum_{r \in R} \left( \frac{\alpha q_A^r + (1 - \alpha) q_B^r}{0} \right) \rho^r (x) \, dx - \sum_{j=1}^n \left( \frac{\alpha q_A^j + (1 - \alpha) q_B^j}{0} \right) \xi^j (x) \, dx
\]

Since we assume that \(\rho^r (x)\) are decreasing and concave, and that \(\xi^j (x)\) are increasing and convex, we have

\[
\frac{\alpha q_A^r + (1 - \alpha) q_B^r}{0} \rho^r (x) \, dx \geq \alpha \int_0^{q_A^j} \rho^r (x) \, dx + (1 - \alpha) \int_0^{q_B^j} \rho^r (x) \, dx
\]

\[
\frac{\alpha q_A^j + (1 - \alpha) q_B^j}{0} \xi^j (x) \, dx \geq \alpha \int_0^{q_A^j} \xi^j (x) \, dx + (1 - \alpha) \int_0^{q_B^j} \xi^j (x) \, dx
\]

\(
\iff \pi^i \left( \{\alpha q_A^r + (1 - \alpha) q_B^r\} \right) \geq \alpha \pi^C \left( \{q_A^r\} \right) + (1 - \alpha) \pi^C \left( \{q_B^r\} \right).
\)

Hence \(\pi^i\) is concave.

QED. ■
**Corollary 8** The optimization problem shown above has a unique optimum.

**Proof.** of Corollary 8:

The objective function is concave, and as long as we define the compact set over which it is optimized, the Corollary is true. Define the compact set by defining the interval for each $q^r$ to be the closed interval $[0, \bar{q}^r]$, where $\bar{q}^r$ is the maximum quantity that can be sold in market $r$. Since the demand function is decreasing and concave, $\bar{q}^r$ is finite. Hence the objective function is optimized over a compact set.

QED. ■

For the centralized problem, there are no marginalizations, thus

$$
\rho^r (q^r) = MR^r (q^r)
$$

$$
\xi^j (q^j) = MC^j (q^j)
$$

With the decentralized problem, for the leader, the effective reduced marginal revenue function for market $r$ is the original marginal revenue function marginalized $\delta (i, r)$ times. The cost function for node $j$ is marginalized $\delta (i, j)$ times. Thus

$$
\rho^r (q^r) = F_{\delta(i,r)} (MR^r (q^r))
$$

$$
\xi^j (q^j) = F_{\delta(i,j)} (MC^j (q^j))
$$

9.2 Proofs for Section 3:

**Proof.** of Lemma 1:

Consider the reduced problem for such networks, as shown in equation 8. This can be expressed in terms of the objective function in Section 9.1 in the following manner

$$
\rho^r (q^r) = MR^r (q^r)
$$

$$
\xi^j (q^j) = MC^j (q^j).
$$

Hence there exists a unique solution. QED. ■

**Proof.** if Lemma 2:

If the condition is violated for any market, it follows that even in the absence of other markets (retailers) the network will not sell in that market. QED. ■

9.3 Proofs for Section 4:

**Proof.** of Lemma 3:

(a) $F_x (f_1 (x) + f_2 (x)) = f_1 (x) + f_2 (x) + x \frac{\partial f_1 (x) + f_2 (x)}{\partial x} = f_1 (x) + x \frac{\partial f_1 (x)}{\partial x} + f_2 (x) + x \frac{\partial f_2 (x)}{\partial x}$

$b) F_x (\alpha f (x)) = \alpha f (x) + x \frac{\partial \alpha f (x)}{\partial x} = \alpha \left( f (x) + x \frac{\partial f (x)}{\partial x} \right) = \alpha F_x (f)$.

QED. ■

**Proof.** Of Lemma 4:

(a) If $\frac{\partial f}{\partial x} \geq 0$ and $\frac{\partial^2 f}{\partial x^2} \geq 0$, then $\frac{\partial}{\partial x} F_x (f (x)) = 2 \frac{\partial f}{\partial x} + x \frac{\partial^2 f}{\partial x^2} \geq 0 \ \forall x > 0$.

(b) If $\frac{\partial f}{\partial x} \leq 0$ and $\frac{\partial^2 f}{\partial x^2} \leq 0$, then $\frac{\partial}{\partial x} F_x (f (x)) = 2 \frac{\partial f}{\partial x} + x \frac{\partial^2 f}{\partial x^2} \leq 0 \ \forall x > 0$.

QED. ■

9.4 Proofs for Section 5:

**Proof.** of Lemma 5:

This follows from the optimization problem. QED. ■

**Proof.** of Lemma 6:
The first-order-condition for \( k \)'s optimization leads to \( w^k = MC^k(q^k) \). Substituting for \( w^k \) in \( j \)'s optimization we obtain

\[
\pi^j (q^j, ..., q^k, ...) = \gamma(.) - MC^k(q^k)q^k.
\]

Thus, instead of two separate problems linked by \( j \) offering \( k \) a wholesale price, the equivalent (single) problem incorporates \( k \)'s optimal decision and \( j \) directly selects \( q^j \). It is also optimal to select \( q^j = q^k \), since otherwise there is either an excess or a shortage, which is unnecessary with deterministic demand.

Note that \( MC^k(q^j)q^j \) is the transformed total cost function for \( j \) (in addition to his original costs which may be present as part of \( \gamma(.) \)). To express this as a (transformed) marginal cost function, we need to find \( \lambda(.) \) such that

\[
\lambda(q^j) = MC^k(q^j) + q^j \frac{\partial}{\partial q^j} MC^k(q^j) = F_{q^j}(MC^k(q^j)).
\]

Hence the equivalent problem for \( j \) after elimination of \( k \) is

\[
\pi^j (q^j, ...) = \gamma(.) - \int_{x=0}^{q^j} F_x(MC^k(x)) dx.
\]

QED. ■

Proof. of Lemma 7:
This follows from the optimization problem. QED. ■

Proof. of Lemma 8:
The first-order-condition for \( k_l \)'s optimization leads to \( w^{k_l} = MR^{k_l}(q^{k_l}) - MC^{k_l}(q^{k_l}) \), for \( l = 1..L \). Without loss of generality, we will show how to incorporate \( k_1 \)'s problem, since all the \( k_l \) are identical.

Eliminating \( w^{k_1} \) we obtain

\[
\pi^j (... , q^{k_1}, ...) = \gamma(.) + \left( MR^{k_1}(q^{k_1}) - MC^{k_1}(q^{k_1}) \right) q^{k_1}.
\]

It is also optimal to select \( q^j = q^k \), since otherwise there is either an excess or a shortage, which is unnecessary with deterministic demand.

Note that in a manner similar to the proof of Lemma 6, \( MC^{k_1}(q^{k_1})q^{k_1} \) is the transformed total cost function for \( j \) (in addition to his original costs which may be present as part of \( \gamma(.) \)), and the appropriate transformed marginal cost function \( \lambda(.) \) is

\[
\lambda(q^j) = F_{q^{k_1}}(MC^{k_1}(q^{k_1})).
\]

In a similar manner, the transformed marginal revenue function \( \mu(.) \) is

\[
\mu(q^{k_1}) = F_{q^{k_1}}(MR^{k_1}(q^{k_1})).
\]
Hence the equivalent problem for $j$ after elimination of $w^k_1$ is

$$
\pi^j(...,q^{k_1},...,)=\gamma(\cdot)+\int_{x=0}^{q^{k_1}}F_x\left(MR^{k_1}(x)\right)dx-\int_{x=0}^{q^{k_1}}F_x\left(MC^{k_1}(x)\right)dx.
$$

Thus, after incorporation of the marginalized functions of all $k_l$, $l=1..L$, we get

$$
\max \quad \pi^j(...,\{q^{k_l}\},...)=\gamma(\cdot)+\sum_{l=1}^{L}\left[\int_{x=0}^{q^{k_l}}F_x\left(MR^{k_l}(x)-MC^{k_l}(x)\right)dx\right].
$$

Finally, we still need the branch point balance equation since $j$ chooses both $q^j$ and all the $q^{k_i}$. This is

$$q^j=\sum_{l=1}^{L}q^{k_l}.$$

(Note that now $j$’s problem has $L+1$ decision variables: $q^j$ and $L$ values of $q^{k_i}$. If we take the first-order-conditions for $j$’s optimization, then there are $L+1$ equations relating these $L+1$ variables.)

QED.

Proof. of Lemma 9:

On each pass through the algorithm, at least one node is eliminated. Since the number of nodes in the network is finite, the algorithm terminates after a finite number of passes.

QED.

Proof. of Lemma 10:

The reduced problem for such networks can be expressed in terms of the objective function in Section 9.1 in the following manner

$$
\rho^r (q^r) = F^{\delta(r,i)}_{q^r}(MR^r (q^r))
$$

$$
\xi^j (q^j) = F^{\delta(i,j)}_{q^j}(MC^j (q^j))
$$

Hence there exists a unique solution. QED.

9.5 Proofs from Section 6:

Proof. of Lemma 11:

This follows from the observations that $q^{1,2}$ is zero for low $w^1$, and increases with increasing $w^1$. Hence, if the constraint is not binding, node 2 can increase his profits by reducing $w^1$ by a small amount.

QED.

Proof. of Lemma 12:

Restricting ourselves to the case where $q^2$ is positive, there are two cases.

In the first case $q^3$ is zero, since $w^1$ is very large and $w^1 = MC^1 (q^{1,2})$. If $q^3 = 0$ then

$$\frac{d q^3}{dw^1} = 0$$

$$\frac{d q^2}{dw^1} = \frac{1}{\frac{d q^2}{dq^2}MC^1 (q^2)}.$$
Since $\frac{dq^2}{dw^1} \neq 0$ for any positive $q^2$, the first-order-condition is equivalent to
\[
\frac{d\pi^2}{dw^1} = \frac{\partial \pi^2}{\partial q^2} \frac{dq^2}{dw^1} + \frac{\partial \pi^2}{\partial q^3} \frac{dq^3}{dw^1}
\]
\[
= \left( -F_{q^3} (MC^1 (q^3)) + (MR^2 (q^2) - MC^2 (q^2)) \right) \frac{dq^2}{dw^1}
\]
\[
= \left( (MR^2 (q^2) - MC^2 (q^2)) - F_{q^2} (MC^1 (q^2)) \right) = 0.
\]

Using the linear marginal functions we get
\[
\frac{dx^2}{dw^1} = \left( (a^2 - 2b^2 q^2) - (a^2 + 2K^2 q^2) - (c^2 + 4K^1 (q^2)) \right) \frac{1}{2K^1} = 0
\]

or $q^2 = \frac{a^2 - c^2 - c_1}{2b^2 + 2K^2 + 4K^1}$

and using $q^2 = q^{1,2*} = \frac{w^1 - c^1}{2K^1}$ from the stage 1 problem when $q^3 = 0$, we get
\[
w^1 = c^1 + 2K^1 q^2
\]
\[
= c^1 + 2K^1 \frac{a^2 - c^2 - c_1}{2b^2 + 2K^2 + 4K^1}.
\]

In the second case where both $q^2$ and $q^3$ are positive, we have
\[
\frac{dq^3}{dw^1} = \frac{1}{\frac{d}{dq^3} (MR^3 (q^3) - MC^3 (q^3))}
\]
\[
\frac{dq^2}{dw^1} = \frac{1}{\frac{d}{dq^3} (MR^3 (q^3) - MC^3 (q^3)) - MC^1 (q^2 + q^3)} \frac{dq^3}{dw^1}
\]

and this simplifies to the following with linear marginal functions
\[
\frac{dq^3}{dw^1} = \frac{1}{\frac{d}{dq^3} (MR^3 (q^3) - MC^3 (q^3))}
\]
\[
= \frac{1}{\frac{d}{dq^3} ((a^3 - 4b^3 q^3) - (c^3 + 4K^3 q^3))}
\]
\[
= \frac{4b^3 + 4K^3}{\frac{d}{dq^3} ((a^3 - 4b^3 q^3) - (c^3 + 4K^3 q^3))}
\]
\[
\frac{dq^2}{dw^1} = \frac{\frac{d}{dq^3} (MR^3 (q^3) - MC^3 (q^3)) - MC^1 (q^2 + q^3)}{\frac{d}{dq^3} MC^1 (q^2 + q^3)} \frac{dq^3}{dw^1}
\]
\[
= \frac{\frac{d}{dq^3} (((a^3 - 4b^3 q^3) - (c^3 + 4K^3 q^3)) - (c^3 + 2K^1 (q^2 + q^3)))}{\frac{d}{dq^3} (c^1 + 2K^1 (q^2 + q^3))}
\]
\[
= - \frac{(4b^3 + 4K^3 + 2K^1)}{2K^1} \frac{1}{4b^3 + 4K^3}
\]
\[
= \frac{2b^3 + 2K^3 + K^1}{K^1 (4b^3 + 4K^3)}.
\]
Observe that

\[ w_D^1 = F(MR^3(q^3) - MC^3(q^3)) \]
\[ = a^3 - c^3 - 4(b^3 + K^3)q^3 \]
\[ = a^3 - c^3 - 4(b^3 + K^3) \frac{a^3 - c^3 - c^3}{4b^3 + 4K^3 + 4K^1} \]
\[ = a^3 - c^3 - 4(b^3 + K^3) \frac{a^3 - c^3}{4b^3 + 4K^3 + 4K^1} - 4(b^3 + K^3) \frac{-c^3}{4b^3 + 4K^3 + 4K^1} \]
\[ = (a^3 - c^3) \left(1 - \frac{4(b^3 + K^3)}{4b^3 + 4K^3 + 4K^1}\right) + \frac{4(b^3 + K^3)c^3}{4b^3 + 4K^3 + 4K^1} \]
\[ = (a^3 - c^3) \left(\frac{4K^1}{4b^3 + 4K^3 + 4K^1}\right) + \frac{b^3 + K^3}{b^3 + K^3 + K^1}c^3. \]

Note that \( q^3 = 0 \) at \( w^1 = a^3 - c^3 \). Also, note that \( \frac{K^1(4b^3 + 4K^3)}{K^1 + 2b^3 + 2K^3} < 2K^1 \), and that \( w^1(q^2) \) is continuous, piecewise linear with a kink at \( a^3 - c^3 = c^3 + 2K^1q^2 \).

QED. ■

**Proof.** of Lemma 13:
Refer to Figure 10. The implied marginal cost function for retailer 2 is given by \( F(w^1(q^2)) \).
When \( f(x) = a + bx \) is linear, the operation \( F_x(f(x)) = F_x(a + bx) = (a + bx) + x.b = a + 2bx \), effectively doubles the slope, keeping the intercept unchanged. However, \( w^1(q^2) \) is differentiable everywhere except at the kink (note that the derivatives from the left and right exist).
Thus, in the segment DB, the marginalized cost function is DE. In the segment BC, which is actually the line AFG, the marginalized cost function is AFG. Thus, the total marginalized cost function is DEFG, with the line segment ending at F is open ended (by the construction of \( F(x) \)).

QED. ■

**Proof.** of Lemma 14:
If we incorporate \( w^1(q^2) \) into retailer 2’s objective function, we obtain

\[ \pi^2(q^2) = -w^1(q^2)q^2 + \int_{0}^{q^2} \{MR^2(x) - MC^2(x)\} dx. \]

The first order condition is

\[ -F_{q^2}(w^1(q^2)) + MR^2(q^2) - MC^2(q^2) = 0. \]

Here \( MR^2(q^2) \) is decreasing and \( MC^2(q^2) \) is increasing in \( q^2 \) by assumption. Finally, from Lemma 13, \( F_{q^2}(w^1(q^2)) \) is increasing. Hence, if

\[ MR^2(0) - MC^2(0) > + F_{q^2}(w^1(q^2))|_{q^2=0} \]

then there is a unique positive \( q^2 \) solution to retailer 2’s problem. Otherwise, the solution is \( q^2 = 0 \).

QED. ■

**Proof.** of Lemma 15:
This follows from the optimization problem. QED. ■

**Proof.** of Lemma 16:
The reduction of the optimization problems for nodes \( k_2..k_L \) is similar to that in the proof of 8.
QED. ■

**Proof.** of Lemma 17:
Consider the first order condition for $j$'s objective function as given in Problem 6 with respect to the quantity sold to a downstream child:

\[
\frac{\partial \pi^j}{\partial q^k} = F_{q^k} \left( MR^k (q^k) - MC^k (q^k) \right) - MC^j (q^j) = 0 \quad k \in C_d (j)
\]

Since these must hold for each child $k$ for which $q^k > 0$, it follows that for any two children $k_a$ and $k_b$ where $k_a, k_b \in C_d (j)$ and in equilibrium $q^{k_a} > 0, q^{k_b} > 0$ we must have

\[
F_{q^{k_a}} \left( MR^{k_a} (q^{k_a}) - MC^{k_a} (q^{k_a}) \right) = F_{q^{k_b}} \left( MR^{k_b} (q^{k_b}) - MC^{k_b} (q^{k_b}) \right).
\]

QED. ■
Appendix 2: Two Part Tariff Contracts

In this section we will show that offering two-part tariffs instead of wholesale prices will leave the follower’s marginal functions unmodified when incorporating them into the parent’s optimization. We will do this for the three procedures that we have used in Sections 5.1.1, 5.1.2 and 6.3.1.

10.1 Manufacturing Leaf Elimination for Two-Part Tariffs

Modify Problem 1 so that \( j \) offers a two-part tariff \((W^{kj}, w^k)\) to a manufacturing leaf \( k \) where \( W^{kj} \) is the side payment from \( k \) to \( j \) and \( w^k \) is the per-unit wholesale price. Define the member problems as given in Problem 7. As before, define the reduced problem for member \( j \) as given in Problem 8.

**Problem 7** The optimization problem with two-part tariffs for \( j \) can be rewritten as

\[
\max_{q^j, \ldots, \{w^k, W^{kj}\}, \ldots} \pi^j \left( q^j, \ldots, \left\{ w^k, W^{kj} \right\}, \ldots \right) = w^j q^j I_d (j) - w^j q^j I_u (j)
\]

\[+ \gamma (.) + W^{kj} - w^k q^k - \int_{x=0}^{q^k} MC^j (x) \, dx,\]

s.t. \( q^j \leq q^k, \Gamma.\)

The optimization problem for \( k \) is

\[
\max_{q^k} \pi^k \left( q^k; W^{kj}, w^k \right) = -W^{kj} + w^k q^k - \int_{x=0}^{q^k} MC^k (x) \, dx.
\]

**Problem 8** Rewrite Problem 7 by replacing \((W^{kj} - w^k q^k)\) with \(-\int_{x=0}^{q^k} MC^k (x) \, dx\) to get the objective function for \( j \) as

\[
\max_{q^j, \ldots} \pi^j \left( q^j, \ldots \right) = w^j q^j I_d (j) - w^j q^j I_u (j)
\]

\[+ \gamma (.) - \int_{x=0}^{q^j} MC^k (x) \, dx - \int_{x=0}^{q^j} MC^j (x) \, dx,\]

s.t. \( \Gamma.\)

**Lemma 19** The optimal solution to \( j \)'s optimization in Problem 7 (incorporating optimal behavior by \( k \)) and the optimal solution to \( j \)'s optimization in Problem 8 have identical \( q^j \) and \( \pi^j \). Thus, if the leader offers a two-part tariff, then the induced transformation is the identity transformation \( I \).

**Proof.** of Lemma 19: Since the side payment \( W^{kj} \) is independent of \( q^k \), the first order condition for \( k \)'s optimization leads to

\[
w^k = MC^k \left( q^k \right),
\]

which is unchanged from Problem 1. Thus, we incorporate this into \( j \)'s optimization and eliminate \( w^j \). Similar to the argument in the Proof of Lemma 6, we can argue that \( q^j = q^k \). Hence we obtain

53
the modified optimization for $j$ as
\[
\max \quad \pi^j (q^j, ..., W^{kj}, ...) = w^j q^j I_d (j) - w^j q^j I_u (j)
\]
\[
+\sum_{k} (\cdot) + W^{kj} - MC^k (q^j) q^j - \int_{x=0}^{q^j} MC^j (x) \, dx,
\]
s.t. $\Gamma$.

Now $j$ can extract all profits that $k$ makes by requiring a side payment of that magnitude (the result stays unchanged if $k$ has a participation constraint, i.e., $k$ must make some exogenously specified minimum profit). Thus, $j$ will set $w^k$ to be the wholesale price that maximizes $\pi^k$ with a wholesale-price-only-contract, i.e.
\[
W^{kj} = \max_{w^k} \pi^k (q^k; 0, w^k).
\]

Now $k$’s profit function with no side payment (i.e., $W^{kj} = 0$ in equation (35)) is
\[
\pi^k (q^k; 0, w^k) = w^k q^k - \int_{x=0}^{q^k} MC^k (x) \, dx.
\]

Substituting for $w^k$ from the first order condition given by equation (37) into $k$’s profit function we get
\[
W^{kj} = MC^k (q^k) q^k - \int_{x=0}^{q^k} MC^k (x) \, dx.
\] (39)

Utilizing $q^k = q^j$, and substituting for $W^{kj}$ from equation (39) into (38) we obtain
\[
\max \quad \pi^j (q^j, ...) = w^j q^j I_d (j) - w^j q^j I_u (j)
\]
\[
+\gamma (\cdot) - \int_{x=0}^{q^j} MC^k (x) \, dx - \int_{x=0}^{q^j} MC^j (x) \, dx,
\]
s.t. $\Gamma$.

QED. ■

10.2 Retailer Leaf Elimination from a Branch Point Downstream of the leader using Two-Part Tariffs

We establish a similar result for the retailer problem by modifying Problems 3 and 4 to consider two-part tariff contracts, where nodes $k$ are the retailer leaves.

**Problem 9** The optimization problem with two part tariffs for the retailer leaves’ parent node $j$ can
be rewritten as

$$\max_{q^i, \ldots, \{w^k, W^kj\}, \ldots} \pi^j \left( q^j, \ldots, \{w^k, W^kj\}, \ldots \right)$$

$$= -w^j q^i I_u (j) + \sum_{k \in C_d (j)} \left( W^kj + w^k q^k \right) - \gamma (.) - \int_{x=0}^{q^j} MC^j (x) \, dx$$

s.t. $q^{j,k} \geq q^k \quad k \in C_d (j)$,

$q^j = \sum_{k \in C_d (j)} q^{j,k}$

$\Gamma$.

The optimization for a retailer $k$ can be rewritten as

$$\max_{q^k} \pi^k \left( q^k; w^k, W^kj \right) = \int_{x=0}^{q^k} MR^k (x) \, dx - W^kj - w^k q^k$$

$$- \int_{x=0}^{q^k} MC^k (x) \, dx$$

Problem 10 Rewrite the objective function for $j$ in Problem 9 by replacing each $(W^kj + w^k q^k)$ with

$$\int_{x=0}^{q^k} MR^k (x) \, dx - \int_{x=0}^{q^k} MC^k (x) \, dx$$

to get

$$\max_{q^i, \{q^k\}_{k \in C_d (j)}, \ldots} \pi^j \left( q^j, \{q^k\}_{k \in C_d (j)}, \ldots \right)$$

$$= -w^j q^i I_u (j) + \sum_{k \in C_d (j)} \left( \int_{x=0}^{q^k} MR^k (x) \, dx - \int_{x=0}^{q^k} MC^k (x) \, dx \right)$$

$$- \gamma (.) - \int_{x=0}^{q^j} MC^j (x) \, dx$$

s.t. $q^j = \sum_{k \in C_d (j)} q^k$

$\Gamma$.

Lemma 20 The optimal solution to $j$’s optimization in Problem 9 (incorporating optimal behavior by the nodes $k$) and the optimal solution to $j$’s optimization in Problem 10 have identical $q^j$ and $\pi^j$. Thus, if the leader offers a two-part tariff, then the induced transformation is the identity transformation $I$.

Proof. of Lemma 20:
The first order condition for $k$’s optimization in Problem 9 leads to

$$w^k = MR^k (q^k) - MC^k (q^k)$$

which is unchanged from Problem 3, the wholesale-price-only contract case.
Eliminating $w^k$ and using $q^j = q^k$, we obtain

$$\max \pi^j \left( q^j, ..., W^{kj}, ... \right) = -w^j q^j I_u \left( j \right)$$

$$+ W^{kj} + \left( MR^k \left( q^j \right) - MC^k \left( q^j \right) \right) q^k - \gamma . - \int_{x=0}^{q^j} MC^j \left( x \right) dx$$

s.t. $\Gamma$.

Note that $j$ can now use $W^{kj}$ to extract all excess profits from $k$, hence

$$W^{kj} = \max_{w^k} \pi^k \left( q^k; 0, w^k \right)$$

or

$$W^{kj} = \max_{w^k} \left[ -w^k q^k + \int_{x=0}^{q^k} MR^k \left( x \right) dx - \int_{x=0}^{q^k} MC^k \left( x \right) dx \right]$$

$$= - \left[ MR^k \left( q^k \right) - MC^k \left( q^k \right) \right] q^k + \int_{x=0}^{q^k} MR^k \left( x \right) dx - \int_{x=0}^{q^k} MC^k \left( x \right) dx$$

Substituting for $W^{kj}$ back into $j$’s objective function (and substituting $q^j = q^k$) we get

$$\max \pi^j \left( q^j, ... \right) = -w^j q^j I_u \left( j \right)$$

$$+ \int_{x=0}^{q^j} MR^k \left( x \right) dx - \int_{x=0}^{q^j} MC^k \left( x \right) dx - \gamma . - \int_{x=0}^{q^j} MC^j \left( x \right) dx$$

s.t. $\Gamma$.

QED. $\blacksquare$

### 10.3 Retailer Leaf Elimination from a Branch Point Upstream of the leader using Two-Part Tariffs

Finally, we consider the branch point upstream of the leader by modifying Problems 5 and 6 to consider two-part tariff contracts, where nodes $k$ are the retailer leaves and node $k_1$ is the leader.

**Problem 11** Node $k_1$’s optimization is

$$\max_{..., w^j, ..., W^{j_{k_1}}, ...} \pi^{k_1} \left( ..., w^j, ..., W^{j_{k_1}}, ... \right) = \gamma . + W^{j_{k_1}} - w^j q^j_{k_1}$$

s.t. $q^{k_1} \leq \ \Gamma$.

The term $\gamma .$ in equation (43) includes $k_1$’s costs and revenues from contracting with children other than $j$. The set $\Gamma$ contains the other constraints from $k_1$’s optimization from contracting with children other than $j$.  

56
The optimization for the branch node \( j \) (which has a downstream parent \( k_1 \)) can be rewritten as

\[
\max_{q^{j,1}, \{q^k, w^k, W^{kj}\}_{k \in C_d(j)}} \pi^j \left( q^{j,1}, \left\{ q^k, w^k, W^{kj}\right\}_{k \in C_d(j)}; w^j, W^{j,k_1} \right) = -W^{j,k_1} + w^j q^{j,1} + \sum_{k \in C_d(j)} \left( W^{kj} + w^k q^k \right) - \int_{x=0}^{q^j} MC^j(x) \, dx
\]

s.t. \( q^j \geq q^{j,k} \quad \forall k \in C_d(j) \)

\[
q^j = q^{j,1} + \sum_{k \in C_d(j)} q^{j,k}.
\]

Since the procedure by which we eliminate nodes first eliminates all upstream children, node \( j \)'s optimization is completely specified here.

The optimization for a retailer leaf \( k \) (which has no children, and has an upstream parent, node \( j \)) can be rewritten as

\[
\max_{q^k} \pi^k \left( q^k; w^k, W^{kj} \right) = \int_{x=0}^{q^k} MR^k(x) \, dx - W^{kj} - w^k q^k - \int_{x=0}^{q^j} MC^j(x) \, dx
\]

Define a modified problem for \( j \) and \( k_1 \) as shown in Problem 6.

**Problem 12** Rewrite \( j \)'s problem in Stage 2:

\[
\max_{q^j, \{q^k\}_{k \in C_d(j)}} \pi^j \left( q^j, \left\{ q^k\right\}_{k \in C_d(j)}; w^j, W^{j,k_1} \right)
\]

\[
= -W^{j,k_1} + w^j q^{j,1} + \sum_{k \in C_d(j)} \left( \int_{x=0}^{q^k} \left\{ MR^k(x) - MC^k(x) \right\} \, dx \right) - \int_{x=0}^{q^j} MC^j(x) \, dx
\]

s.t. \( q^j = q^{j,1} + \sum_{k \in C_d(j)} q^k \bigg\} \Gamma. \]

We have \( k_1 \)'s problem (unchanged) in Stage 1:

\[
\max_{\ldots, \{w^j, W^{j,k_1}\}, \ldots} \pi^{k_1} \left( \ldots, \left\{ w^j, W^{j,k_1}\right\}, \ldots \right) = \gamma(.) + W^{j,k_1} - w^j q^{j,k_1}
\]

s.t. \( q^{k_1} \leq q^{j,k_1} \bigg\} \Gamma.

**Lemma 21** The optimal solution to \( j \)'s optimization in Problem 11 (incorporating optimal behavior by \( k \)) and the optimal solution to \( j \)'s optimization in Problem 12 have identical \( q^j, \{q^k\}_{k \in C_d(j)} \) and \( \pi^j \).
Proof. of Lemma 21:
This is similar to the proof of Lemma 20 since we are eliminating retailer nodes downstream of the
branch point using two-part tariffs.
QED. □

We have established that all the marginal functions for the children nodes of \( j \) are incorporated
into \( j \)'s optimization without any transformation. We still have the task of establishing that when
\( k_1 \) offers a two-part tariff to \( j \), then \( j \)'s marginal functions can be incorporated into \( k_1 \)'s optimization
without any transformation. However, looking at Problem 12 we see that it has a structure very
similar to the manufacturing node elimination problem, Problem 7. Hence, with two part tariffs, \( j \)'s
marginal functions are incorporated unchanged into \( k_1 \)'s problem from Lemma 19.

Define the modified problem for \( k_1 \) as shown in Problem 13.

Problem 13 Rewrite \( k_1 \)'s problem as

\[
\max_{q^{k_1}, q^j, \{q^k\}_{k \in C_d(j)}} \pi^{k_1} \left( q^{k_1}, q^j, \{q^k\}_{k \in C_d(j)} \right) \quad (48)
\]

\[
= \gamma(.) + \sum_{k \in C_d(j)} \left( \int_{x=0}^{q^k} \left\{ MR^k(x) - MC^k(x) \right\} dx \right) - \int_{x=0}^{q^j} MC^j(x) dx \quad (49)
\]

s.t. \( q^j = q^{k_1} + \sum_{k \in C_d(j)} q^k \)

\( \Gamma \).

Observe that earlier, in problems 11 and 12, we had shown only the part of \( k_1 \)'s optimization
which was relevant to the contract between \( k_1 \) and \( j \). The variable \( q^{k_1} \) was not included in these two
problems. However, with the second reduction, we have the property that \( q^{j,k_1} = q^{k_1} \). Hence in
Problem 13, the constraint (49) contains the variable \( q^{k_1} \).

Lemma 22 The optimal solution to \( k_1 \)'s optimization in Problem 12 (incorporating optimal behavior
by \( j \)) and the optimal solution to \( k_1 \)'s optimization in Problem 13 have identical \( q^{k_1}, q^j, \{q^k\}_{k \in C_d(j)} \)
and \( \pi^{k_1} \).

Proof. of Lemma 22:
This is similar to the proof of Lemma 19 since we are eliminating an upstream node using two-
part tariffs. While in Lemma 19 the upstream node had only a marginal cost function, and here the
upstream node has both marginal costs and marginal revenues, the arguments are analogous.
QED. □