Decentralized Pricing and Capacity Decisions in a Multi-Tier System with Modular Assembly

Fernando Bernstein       Gregory A. DeCroix

The Fuqua School of Business, Duke University, Durham, NC 27708

December 2002
Revised September 2003

Abstract

We model a modular assembly system in which a final assembler outsources some of the assembly task to first-tier suppliers (subassemblers), who produce modules made up of multiple components. The assembler sets module prices it will pay to the subassemblers, the subassemblers set component prices they will pay to suppliers, and then all players choose how much capacity to install, with the minimum capacity choice determining system capacity. Finally, stochastic end-product demand is observed and all players produce (and are paid for) the same number of units – the minimum of demand and system capacity.

We characterize equilibrium price and capacity choices, and then use that characterization to derive results regarding higher-level structural choices by the assembler – such as how to group components into modules and which suppliers to choose as subassemblers. We also compare performance of the system to a traditional assembly system with an assembler and suppliers but without subassemblers.
1 Introduction

In recent years many manufacturing companies have started shifting from a traditional final-assembly approach to an approach often referred to as modular assembly. Under the traditional approach, the manufacturer purchases components from its first-tier suppliers, and then assembles those components to produce the finished product. Under modular assembly, the finished product is divided into modules, and a relatively small number of first-tier suppliers are chosen as subassemblers. Each subassembler is responsible for delivering a module to the manufacturer – i.e., the subassembler purchases the components contained in the module (from a subset of the former first-tier suppliers) and assembles those components to produce the module. The manufacturer then assembles the modules to produce the finished product. The manufacturer is essentially outsourcing some of the assembly work to a subset of its first-tier suppliers.

Modular assembly is becoming a factor in a variety of industries, but it has received particular attention in the automobile industry. It has been referred to as the “Automotive New World Order” (Green 1998), and “a trend that is more revolutionary than the moving assembly line” (Slaughter 1999). Several major manufacturers have instituted modular assembly or are considering doing so. For example, Volkswagen has implemented an extreme version of modular assembly at its truck and bus plant in Resende, Brazil. The vehicles are split into seven modules (e.g., cabin, chassis, engine/transmission, suspension/axles/wheels, etc.), which are assembled and installed on the finished product by seven suppliers; see e.g., Lippert (1998) and Smith (2002). Nissan has implemented modular assembly at its new plant in Mississippi, where a few suppliers deliver assembled vehicle sections. This is in contrast to Nissan’s 20-year-old plant in Smyrna, Tennessee, where most of the assembly is done at the plant; see Chappell (2001). General Motors, Ford and DaimlerChrysler have also implemented modular assembly.

The primary factor motivating manufacturers to make this shift is a search for cost reductions. By shifting responsibility for a portion of assembly to their suppliers, manufacturers hope also to shift costs associated with design and engineering, supplier coordination, inventory, assembly capacity and labor. In some cases suppliers have lower assembly-related costs than finished-goods manufacturers, so this shift is not merely a transfer of costs between members of the supply chain, but actually reduces total supply chain costs. One reason for this cost advantage is that employees of the manufacturer may be unionized, whereas employees of suppliers may not be. For example, non-unionized employees of Volkswagen’s suppliers in Resende, Brazil, are reported to earn approximately one third the wage of unionized em-
ployees of Volkswagen, Ford or Mercedes in São Paulo (Slaughter 1999). Also, workers at the Big Three plants in the U.S. make around $20 an hour, compared to $13 or less an hour for workers in supplier plants (Welch 2001).

Outsourcing subassembly work has some risks, however. By giving suppliers more assembly responsibilities, the firm also gives them more influence in the supply chain. These subassemblers may be able to use this influence to capture higher profit margins at the expense of the manufacturing firm. The firm’s ties to its former first-tier suppliers are weaker, and it can only influence their actions indirectly (through the subassembler), which may result in a further loss of control. Other risks/challenges the assembler faces with modular assembly include loss of knowledge about supplier costs and technological capabilities, possible unionization of (now higher-profile) non-union subassembly facilities, potential quality problems in module assembly, and the proper way to assign responsibility for resulting warranty costs. (Novak and Tayur 2003 examine some aspects of the warranty-cost issue in modular assembly systems.)

Our analysis focuses on the trade-off between cost reductions and loss of control that an assembler faces when shifting to modular assembly. The objective of this paper is to develop a better understanding of how such a shift will affect various aspects of supply chain performance. With a better understanding of both the positive and negative impacts, firms can then make better-informed decisions regarding a move to modular assembly.

We begin with an in-depth analysis of a multi-tier assembly system representing a modular assembly environment. For a given configuration (i.e., a fixed set of modules and a fixed set of subassemblers) we partially characterize the optimal prices for the assembler to offer the subassemblers, the equilibrium component prices the subassemblers set in response, and the equilibrium (component and assembly) capacities that will result. Based on these results, we examine higher-level questions about the design of the modular assembly system. These questions include: Should the system be balanced (i.e., modules of similar size or cost) or unbalanced? Which suppliers should be chosen as subassemblers? How much assembly work should be outsourced? We also show that, if the subassemblers do not have lower assembly-related costs than the manufacturer, then modular assembly results in strictly lower capacity and strictly lower expected profit for both the manufacturer and the system when compared to traditional assembly.

If subassemblers do have lower assembly-related costs, these lower costs may (or may not) counter the profit loss from modular assembly. In particular, we show that when the manufacturer outsources a sufficiently large portion of assembly, it is more likely that the cost
savings outweigh the profit loss, making modular assembly attractive to the manufacturer. In addition, we perform numerical studies that suggest that profitable modular assembly is also more likely when assembly costs represent a large portion of total system costs, and is less likely when demand variance is high.

Several recent papers study settings that share some features with our model. Lariviere and Porteus (2001) analyze the problem of a single supplier who must choose the wholesale price at which to sell a product to a single retailer. The retailer then decides how many units to purchase based on that price, the per unit revenue and a demand forecast. The paper identifies conditions on the demand distribution that ensure that the supplier’s profit is unimodal as a function of the wholesale price. Cachon and Lariviere (2001) consider a setting where a downstream manufacturer chooses contract terms, as in the current paper, but where there is just a single supplier of a critical component. The manufacturer presents the supplier with a demand forecast and a contract offer. The supplier then builds capacity based on the initial demand forecast and contract terms, after which the manufacturer observes actual demand and submits a final order. The paper explores forecast sharing by the manufacturer and the capacity decision by the supplier under two different compliance regimes – forced compliance and voluntary compliance. Under voluntary compliance, the supplier chooses the capacity that maximizes her profit.

The literature most relevant to this paper is research studying the behavior of decentralized assembly systems. Gurnani and Gerchak (1998) consider an assembly system where production yield is uncertain for each of the components but end-product demand is deterministic. They analyze the performance of different contract structures and identify two types of contracts that allow the system to match the performance of a centralized system. Gerchak and Wang (2002) study a decentralized assembly system with perfect component yield and stochastic end-product demand, where firms make a capacity choice prior to the realization of demand. They examine the behavior of the system under supplier pricing and assembler pricing. They characterize the optimal prices and resulting equilibrium production quantities for that system. Similar results are obtained in Tomlin (2000, 2001). Wang and Gerchak (2003) generalize Gerchak and Wang (2002) to include a secondary source for components and a salvage value for leftover capacity. They characterize the optimal/equilibrium prices and resulting capacity choices for the system. They also identify a capacity-subsidy contract that can coordinate the channel under certain conditions. All of these papers restrict attention to what we call a traditional assembly system – they do not address multi-tier systems or the issue of modular assembly.
Reviews of the broader literature on decentralized decision making in supply chains can be found in Tsay et al. (1999), Lariviere (1999) and Cachon (2002).

Since modular assembly corresponds to outsourcing part of the final assembly task, the current paper is related to the industrial organization literature on make vs. buy and the boundaries of the firm. This literature can be traced to Coase (1937), who argues that if transactions between firms are costly, then it may be more efficient to conduct certain activities within a single firm. The study of transaction costs continued with Williamson (1975, 1979, 1985), who identifies three transaction characteristics that are important in determining firm boundaries: frequency, uncertainty and asset specificity. A higher level of uncertainty increases the likelihood that future situations will arise that cannot be anticipated at the time a contract is negotiated between two firms and initial investments are made, and therefore adjustments will need to be negotiated at a later date. When the values of initial investments are specific to this particular relationship, one firm may be able to hold-up the other – i.e., capture some of the value of the other firm’s investment. This leads to inefficiency in the form of underinvestment, spending on protection against being held up, etc. The property rights theory pioneered by Grossman and Hart (1986) takes a somewhat different approach to the issues of incomplete contracts and potential hold-ups. In this theory, ownership of non-human assets gives a firm bargaining power when returns are allocated after investments are made. This ex post bargaining power in turn influences the level of ex ante investment the parties are willing to make. (For a summary of the key issues in this literature, as well as additional references, see Holmstrom and Roberts 1998.)

A few other papers address boundary-of-the-firm issues in settings with some features in common with our model. Walker and Weber (1984) provide a statistical analysis of the factors influencing the make-or-buy decision in the U.S. auto industry, finding empirical evidence that supplier cost advantages play a significant role. Lewis and Sappington (1991) explore how the boundary of the firm is influenced by technological change when suppliers have lower costs but are harder to monitor than internal production. Plambeck and Taylor (2002) study the impact on innovation effort when a manufacturer sells its entire production facility to a contract manufacturer. Baker and Hubbard (2002) analyze the interaction between job design and two specific types of technology investments in the trucking industry – one that facilitates greater efficiency for a supplier’s fleet, and one that allows for easier monitoring – on the choice between using a private fleet or contracting for trucking services. McMillan (1990) documents differences between U.S. and Japanese procurement practices and identifies possible theoretical explanations for these practices.
Our setting does involve relationship-specific investments (capacity) and uncertainty (demand), but our model has a different emphasis than those examined in the preceding literature. We analyze a one-period model in which, at the time decisions are made, there is complete clarity regarding who makes what decisions and how payoffs will be determined. So the issues of unanticipated occurrences, ex post bargaining and hold-up that are central to the preceding literature are not a factor here. Instead, we focus somewhat more detailed attention on aspects of the product, cost and supply chain structure. As a result, our perspective is more similar to the supply chain management literature cited above – in which responsibility for pricing and quantity decisions and the presence of double marginalization explain differences in performance between different supply chains. Our results can be seen as complementary to those in the industrial organization literature, however, in the following sense. We explore factors affecting whether assembly outsourcing leads to better or worse performance in the absence of unanticipated occurrences and the need for ex post bargaining. When these factors are present (as would typically be the case in reality), they may add to the arguments against modular assembly that we identify. However, since the total number of inter-firm relationships are the same in both traditional and modular assembly systems (in our model), it is not clear which system would be affected more by renegotiation or other transaction costs. (Note that this differs from the traditional make-or-buy setting, where internal production eliminates the need for inter-firm transactions.) Detailed analysis of this question is beyond the scope of this paper.

The remainder of the paper is organized as follows. Section 2 presents a basic model of the modular assembly setting. In Section 3 we analyze the behavior of the modular assembly system, and in Section 4 we use these results to answer questions regarding design of the system. Section 5 addresses the profitability of modular assembly in the presence of assembly cost savings, and Section 6 provides some concluding remarks. All proofs are relegated to an Appendix.

2 Model

We consider a decentralized multi-tier modular assembly system that produces a single product to meet uncertain demand during a single selling season. The system consists of a single final assembly firm, several (first-tier) subassemblers, and several (second-tier) suppliers. Each supplier produces a single component (or kit of components), which he then ships to one of the subassemblers. Each subassembler assembles the components she receives (along
with a component she produces herself) into a module, which she then ships to the assembler. The assembler then assembles these modules into units of the finished product, which he uses to meet end customer demand. This contrasts with the traditional assembly system in which there are no subassemblers – the assembler simply purchases components from first-tier suppliers and assembles these into the finished product.

Of course, even traditional assembly systems have some degree of modular assembly. The components that first-tier suppliers sell to final assemblers are in fact often made up of many smaller parts from second-tier suppliers, and so on. The issue here is one of degree – the current movement towards modular assembly pushes this modularization much further than traditional assembly systems. Our somewhat simplified representations of modular and traditional assembly systems capture the key differences between these two approaches.

Figure 1 illustrates the distinction between the traditional and modular assembly systems for a given product.

![Figure 1: Traditional and Modular Assembly Systems](image)

The final product consists of $n$ modules, and each module $i$ consist of $m_i + 1$ different parts or components – $m_i$ from the $m_i$ suppliers, plus one from the subassembler. Subassembler $i$ assembles module $i$, while supplier $ij$ supplies component $j$ for module $i$ (where $j = 0$ corresponds to the subassembler’s component). Without loss of generality, one unit of each part is required for each module, and one unit of each module is required for the final product. For now, we assume a specific configuration of modules that leads to the final product and a specific configuration of each set of parts necessary for each module. (We will explore the choice among different system configurations in later sections.)

We model decision making during a single period. Since in many industries the lead times for substantially increasing physical facility capacities are very long – often on the order of 6
months to a year or more – we assume that the capacity decisions $Q_0$ for the assembler, $Q_i$ for subassembler $i$, and $Q_{ij}$ for supplier $ij$ must be made at the beginning of the period. (In fact, the reader can interpret the length of the period to be equal to the lead time for installing additional capacity.) In the model, contracts between the players (i.e., the prices offered by the assembler and subassemblers) last for the duration of the period. In practice, supply contracts vary in length, and in some cases may be longer than capacity lead times – e.g., contracts may last for the lifetime of a given model of the finished product. Modeling such cases explicitly would require multiple periods and multiple capacity-expansion decisions over the course of the contract, adding significant complexity to the model. Our one-period model preserves analytical tractability, but still captures the key elements affecting the choice between modular and traditional assembly. (In the concluding section we discuss how our results may extend to variations of our model.)

Let $D$ represent uncertain demand for the product during the period, with cdf $F$, pdf $f$ and $\overline{F} = 1 - F$. In satisfying demand during the period the system operates as assemble-to-order, i.e., each demand for a unit triggers procurement and assembly operations throughout the chain with negligible delay. This is a reasonable approximation if production lead times are short relative to the selling season – for example, if all firms are geographically close and lead times for assembly are relatively short, as is the case in many assembly systems. (See, e.g., Slaughter 1999, Weber 2003.)

All firms incur costs for each unit of capacity installed. (If the capacity remains useful at the end of the period, these costs can be interpreted as opportunity costs for reserving the capacity for that period.) The per unit capacity costs are given by $k_0$ for assembly operations at the assembler, $k_1, \ldots, k_n$ for assembly operations at the subassemblers, $k_{10}, \ldots, k_{n0}$ for component production operations at the subassemblers, and $k_{i1}, \ldots, k_{imi}$, $i = 1, \ldots, n$ for component production operations at the suppliers. In addition, there are unit component production costs $c_{ij}$ for each supplier $ij$, and $c_{i0}$ for each subassembler $i$. Finally, there are unit assembly costs $c_i$, $i = 1, \ldots, n$ for the subassemblers and $c_0$ for the assembler. Without loss of generality, we assume that the revenue the assembler receives for each unit sold is scaled to one, and we assume that $\sum_{i=1}^{n} \sum_{j=0}^{m_i} (k_{ij} + c_{ij}) + \sum_{i=0}^{n} (k_i + c_i) < 1$ – i.e., the system is profitable. We also assume that the demand distribution and all cost parameters are common knowledge.

Since in practice the shift to modular assembly is generally being driven by final assemblers, we view the system from that perspective – the assembler acts as the leader in decision making. The sequence of events is as follows. First the assembler chooses the prices $\alpha_i$ he
will pay to subassembler $i$ for each unit of module $i$ produced. Then each subassembler $i$ simultaneously chooses the fraction $\gamma_{ij}$ of this to pay supplier $ij$ for each unit of component $ij$ produced – i.e., supplier $ij$ receives $\alpha_i \gamma_{ij}$ per unit produced. Next all players simultaneously install or reserve their production and assembly capacities – the capacity of the overall system is the minimum of these capacities. (Each subassembler will always choose the same capacity for assembly of her module and production of her own component.) Finally, demand is observed, each player produces and/or assembles a quantity equal to the minimum of demand and the overall system capacity, and each player incurs its costs and revenues. For notational convenience, define $\alpha_0 = 1 - \sum_{i=1}^{n} \alpha_i$ to be the net revenue kept by the assembler and $\gamma_{i0} = 1 - \sum_{j=1}^{m_i} \gamma_{ij}$ to be the fraction of the revenue $\alpha_i$ kept by subassembler $i$.

A few aspects of the model deserve a little further discussion. We do not explicitly model alternative (competing) suppliers/subassemblers – instead, assume that the assembler has already selected the suppliers/subassemblers with the lowest cost profiles (i.e., those that will provide the greatest capacity given any unit price). Also, the assembler in our model does not include an explicit capacity clause in the contract he offers the subassemblers. Instead, capacity installation by the subassemblers (and in turn by the suppliers) must be induced by offering sufficiently high prices – i.e., the subassemblers and suppliers control their own capacity choices (once prices are set). This is an approximation of a feature of modular assembly systems that has been observed in practice – that the increased burden placed on subassemblers (i.e., taking on assembly capacity costs) is explicitly linked with increased subassembler influence over system capacity (e.g., see Weber 2003). This corresponds to what Cachon and Lariviere (2001) call voluntary compliance, which arises when enforcement of capacity requirements is difficult. They argue that this is often the case, since the existence of many factors affecting capacity can make it hard for a court to distinguish between under-capacity that is intentional and that which is the result of factors beyond the supplier’s control. They show that in a voluntary compliance setting the addition of a capacity requirement to the contract provides no advantage to the customer (the assembler in our model).

Below is a summary of notation.

\[
\begin{align*}
\alpha_i & = \text{assembler’s per-unit payment (price) to subassembler } i, \ i = 1, \ldots, n; \\
\alpha_0 & = \text{assembler’s per-unit net revenue;} \\
\alpha & = (\alpha_1, \ldots, \alpha_n), \text{ the vector of prices offered by the assembler;} \\
k_0 & = \text{assembler’s unit capacity cost;} \\
k_i & = \text{subassembler } i\text{’s unit capacity cost for assembly, } i = 1, \ldots, n;
\end{align*}
\]
\(k_{i0} = \) subassembler \(i\)'s unit capacity cost for component production, \(i = 1, \ldots, n;\)
\(k_{ij} = \) supplier \(ij\)'s unit capacity cost, \(i = 1, \ldots, n, j = 1, \ldots, m_i;\)
\(c_0 = \) assembler's unit assembly cost;
\(c_i = \) subassembler \(i\)'s unit assembly cost, \(i = 1, \ldots, n;\)
\(c_{i0} = \) subassembler \(i\)'s unit cost for component production, \(i = 1, \ldots, n;\)
\(c_{ij} = \) supplier \(ij\)'s unit cost for component production, \(i = 1, \ldots, n, j = 1, \ldots, m_i;\)
\(\gamma_{ij} = \) fraction of \(\alpha_i\) paid by subassembler \(i\) to supplier \(ij, i = 1, \ldots, n, j = 1, \ldots, m_i;\)
\(\gamma_{i0} = \) fraction of \(\alpha_i\) kept by subassembler \(i, i = 1, \ldots, n;\)
\(\gamma_i = (\gamma_{i1}, \ldots, \gamma_{im_i}), the vector of price shares offered by subassembler \(i, i = 1, \ldots, n;\)
\(\gamma = (\gamma_1, \ldots, \gamma_n), the vector of price shares offered by all subassemblers;\)
\(\gamma_{-i} = \) the vector \(\gamma\) with the components in \(\gamma_i\) removed;
\(Q_0 = \) capacity chosen by the assembler;
\(Q_i = \) capacity chosen by subassembler \(i, i = 1, \ldots, n;\)
\(Q_{ij} = \) capacity chosen by supplier \(ij, i = 1, \ldots, n, j = 1, \ldots, m_i;\)
\(Q = ((Q_{ij}), (Q_i), Q_0), the vector of capacities;\)
\(Q_{-h} = \) the vector \(Q\) with the component corresponding to player \(h\) removed.

3 System Behavior

In this section we analyze the behavior of the system by considering the events in the two-stage game in reverse order. We first investigate the second-stage or capacity game, in which firms simultaneously select capacity levels \(Q\), given a set of prices \(\alpha\) and price shares \(\gamma\). After that we explore the first-stage or pricing game in two steps. First, for a given set of per-unit payments \(\alpha\) by the assembler, we examine the pricing game in which the subassemblers simultaneously choose the price fractions \(\gamma\) to share with their suppliers, anticipating the resulting equilibrium capacity levels \(Q\). Second, we explore the optimization problem in which the assembler selects the vector of payments \(\alpha\) that maximizes his profit, in anticipation of the resulting equilibrium subassembler prices \(\gamma\) and the equilibrium \(Q\) in the second stage capacity game.

3.1 Capacity Game

In the capacity game, for any fixed price vector \(\alpha\) and price share vector \(\gamma\), each player chooses the capacity that maximizes its expected profit, given the capacities chosen by the other players.
For any vector $Q$ of capacities, define

$$Q_{\min}(Q) = \min \left\{ \min_{1 \leq i \leq n, 1 \leq j \leq m_i} \{Q_{ij}\}, \min_{1 \leq i \leq n} \{Q_i\}, Q_0 \right\}. $$

The players face the following capacity-choice problems.

Supplier $ij$: $\max_{Q_{ij}} \pi_{ij}(Q_{ij}|Q_{-ij}, \alpha, \gamma) = -k_{ij}Q_{ij} + (\alpha_i \gamma_{ij} - c_{ij})E \left[ \min(Q_{\min}(Q), D) \right]$;

Subassembler $i$: $\max_{Q_i} \pi_i(Q_i|Q_{-i}, \alpha, \gamma) = -(k_{i0} + k_i)Q_i + (\alpha_i \gamma_{i0} - (c_{i0} + c_i))E \left[ \min(Q_{\min}(Q), D) \right]$;

Assembler: $\max_{Q_0} \pi_0(Q_0|Q_{-0}, \alpha, \gamma) = -k_0Q_0 + (\alpha_0 - c_0)E \left[ \min(Q_{\min}(Q), D) \right]$.

For any player $h$, let $Q^\min_{-h}(Q)$ denote the value of $Q_{\min}(Q)$ with party $h$ excluded from the minimization. It is easy to see that $\pi_h$ is decreasing on $Q_h > Q^\min_{-h}(Q)$, while $Q_{\min}(Q) = Q_h$ on $Q_h \leq Q^\min_{-h}(Q)$, so the above problems can be rewritten as follows:

Supplier $ij$: $\max_{Q_{ij} \leq Q^\min_{-ij}(Q)} \pi_{ij}(Q_{ij}|Q_{-ij}, \alpha, \gamma) = -k_{ij}Q_{ij} + (\alpha_i \gamma_{ij} - c_{ij})E \left[ \min(Q_{ij}, D) \right]$;

Subassembler $i$: $\max_{Q_i \leq Q^\min_{-i}(Q)} \pi_i(Q_i|Q_{-i}, \alpha, \gamma) = -(k_{i0} + k_i)Q_i + (\alpha_i \gamma_{i0} - (c_{i0} + c_i))E \left[ \min(Q_i, D) \right]$;

Assembler: $\max_{Q_0 \leq Q^\min_{-0}(Q)} \pi_0(Q_0|Q_{-0}, \alpha, \gamma) = -k_0Q_0 + (\alpha_0 - c_0)E \left[ \min(Q_0, D) \right]$.

For each player $h$, it is convenient to define that player’s isolated capacity problem. Each isolated capacity problem represents the problem faced by that player if he were allowed to choose production capacity for the entire system. Player $h$’s isolated capacity problem is obtained by dropping the constraint $Q_h \leq Q^\min_{-h}(Q)$ above. Let $Q^I_h$ be the optimal capacity for player $h$ in player $h$’s isolated capacity problem. Each isolated capacity problem is simply a newsvendor problem, so

$$Q^I_{ij}(\alpha, \gamma_{ij}) = \mathcal{F}^{-1}\left( \frac{k_{ij}}{\alpha_i \gamma_{ij} - c_{ij}} \right), \quad Q^I_i(\alpha, \gamma_i) = \mathcal{F}^{-1}\left( \frac{k_{i0} + k_i}{\alpha_i \gamma_{i0} - (c_{i0} + c_i)} \right), $$

and

$$Q^I_0(\alpha_0) = \mathcal{F}^{-1}\left( \frac{k_0}{\alpha_0 - c_0} \right).$$

Observe that $Q^I_{ij}$ is increasing in $\gamma_{ij}$ and $Q^I_i$ is decreasing in $\gamma_i$ since $\gamma_{i0} = 1 - \sum_{j=1}^{m_i} \gamma_{ij}$. Let

$$Q^*_d(\alpha, \gamma) = \min \left\{ \min_{1 \leq i \leq n, 1 \leq j \leq m_i} \{Q^I_{ij}(\alpha, \gamma_{ij})\}, \min_{1 \leq i \leq n} \{Q^I_i(\alpha, \gamma_i)\}, Q^I_0(\alpha_0) \right\}.$$

**Proposition 1** Fix the vectors $\alpha$ and $\gamma$. A capacity vector $Q$ is a Nash equilibrium for the capacity game if and only if $Q_{ij} = Q_i = Q_0 \leq Q^*_d(\alpha, \gamma)$ for all $i$ and $j$. The equilibrium with all capacities equal to $Q^*_d$ is Pareto-optimal, and this is the only Pareto-optimal equilibrium.
3.2 Subassemblers’ Pricing Game

When analyzing the pricing game we assume that, for any choices of $\alpha$ and $\gamma$, all players will choose the Pareto-optimal equilibrium capacity level $Q^*_d(\alpha, \gamma)$ in the subsequent capacity game. As a result, each player’s profit function in the pricing game depends only on $\alpha$ and $\gamma$.

First, suppose that the assembler has already set the vector of prices $\alpha$ that he will pay to the subassemblers. We are interested in the existence and properties of Nash equilibria in the subassemblers’ pricing game that follows. In order for it to be economical for subassembler $i$ to remain in the assembly system, we must have

$$\alpha_i > K^i + C^i \overset{\text{def}}{=} \alpha^\text{min}_i,$$

(1)

where $K^i \overset{\text{def}}{=} \sum_{j=0}^{m_i} k_{ij} + k_i$ and $C^i \overset{\text{def}}{=} \sum_{j=0}^{m_i} c_{ij} + c_i$. We therefore assume that the $\alpha_i$satisfy (1) for all $i = 1, ..., n$.

Given $\alpha$ and $\gamma_{-i}$, subassembler $i$’s pricing problem is as follows:

$$\max_{\gamma_i} \pi_i(\gamma_i|\gamma_{-i}, \alpha) = -(k_{i0} + k_i)Q^*_d(\alpha, \gamma) + (\alpha_i\gamma_{i0} - (c_{i0} + c_i))E[\min(Q^*_d(\alpha, \gamma), D)].$$

The following lemma establishes some structural properties of the optimal solution to this problem. Its proof is immediate from Lemma 1 in Wang and Gerchak (2003). (A similar result appears in Tomlin 2000.)

**Lemma 1** For any $\alpha$ and $\gamma_{-i}$, subassembler $i$’s best response pricing vector $\gamma_i$ satisfies

$$\frac{k_{i1}}{\alpha_i\gamma_{i1} - c_{i1}} = \cdots = \frac{k_{im_i}}{\alpha_i\gamma_{im_i} - c_{m_i}} \geq \frac{k_{i0} + k_i}{\alpha_i\gamma_{i0} - (c_{i0} + c_i)},$$

or equivalently, $Q^I_{i1}(\alpha_i, \gamma_{i1}) = \cdots = Q^I_{im_i}(\alpha_i, \gamma_{im_i}) \leq Q^I_{i1}(\alpha_i, \gamma_i)$.

The equalities in (2) imply that subassembler $i$’s best response vector $\gamma_i$ to $\alpha$ and $\gamma_{-i}$ can be completely characterized by any single price share $\gamma_{ij}, j = 1, ..., m_i$, say $\gamma_{i1}$. This allows us to reduce subassembler $i$’s pricing decision to the one-dimensional problem of choosing $\gamma_{i1}$, with the other price shares $\gamma_{ij}$ given by

$$\gamma_{ij} = \left(\frac{1}{\alpha_i}\right) \left(c_{ij} + \frac{k_{i1}(\alpha_i\gamma_{i1} - c_{i1})}{k_{i1}}\right).$$

(3)

We can also establish bounds on the best response $\gamma_{i1}$. To that end, define

$$\gamma_{1} = \frac{c_{i1} + k_{i1}}{\alpha_i} \quad \text{and} \quad \gamma_{i1} = \frac{k_{i1}}{K^i} + \frac{c_{i1}}{\alpha_i} - \frac{k_{i1}C^i}{\alpha_iK^i}.$$
Since $\gamma_i$ is the minimum possible fraction that allows supplier $i1$ to remain in the system, it follows that $\gamma_i \geq \gamma_i$. Also, it is easy to check that (1) implies $\gamma_i > \gamma_i$, and that (2) and (3) imply $\gamma_i \leq \gamma_i$, yielding

$$\underline{\gamma_i} \leq \gamma_i \leq \overline{\gamma_i}. \tag{5}$$

As a result, in the subassembler pricing game each subassembler $i$ simultaneously chooses the price share $\gamma_i$ satisfying (5) (which in turn determines the remaining $\gamma_{ij}$ via (3)) that maximizes her profit, given the price shares chosen by the other subassemblers. Unfortunately this formulation does not have a structure that lends itself to easy analysis. However, we can transform the game into an equivalent game – in which the subassembler pricing choices are expressed in terms of capacities – that does have a nice structure.

To see this relationship between $\gamma_i$ and module $i$ capacity note that, given subassembler $i$’s best response pricing vector $\gamma_i$, Lemma 1 implies that $F^{-1}\left(\frac{k_i}{\alpha_i \gamma_i - c_i}\right) = Q_i(\alpha_i, \gamma_i) = \min \{ \min_{1 \leq j \leq m_i} \{ Q_{ij}(\alpha_i, \gamma_{ij}) \}, Q_i(\alpha_i, \gamma_i) \}$, where this last quantity can be thought of as the capacity for module $i$ resulting from subassembler $i$’s pricing decision. In other words, there is a one-to-one correspondence between any $\gamma_i$ satisfying Lemma 1 and a capacity choice $Q_i = F^{-1}\left(\frac{k_i}{\alpha_i \gamma_i - c_i}\right)$ for module $i$. As a result, instead of expressing subassembler $i$’s best-response pricing problem in terms of the price share $\gamma_i$ we can express it in terms of the quantity $Q_i$, with the understanding that for any choice of $Q_i$, subassembler $i$ would choose price shares defined by

$$\gamma_i = \frac{c_i}{\alpha_i} + \frac{k_i}{\alpha_i F(Q_i)} \tag{6}$$

and (3). By combining (6) and (5) we obtain the capacity-choice bounds

$$0 \leq Q_i \leq Q_i(\alpha_i) \overset{\text{def}}{=} F^{-1}\left(\frac{K_i}{\alpha_i - C_i}\right). \tag{7}$$

Combining (6) and (3) yields

$$\gamma_{i0} = 1 - \frac{\sum_{j=1}^{m_i} c_{ij}}{\alpha_i} - \frac{\sum_{j=1}^{m_i} k_{ij}}{\alpha_i F(Q_i)}. \tag{8}$$

Finally, if each subassembler $j$ chooses capacity $Q_j$ in the subassembler pricing game, then the Pareto-optimal capacity equilibrium in the second stage capacity game is

$$Q^*_{\alpha}(\alpha, \gamma) = \min \left\{ \min_j Q_j, Q^*_{0}(\alpha_0) \right\}. \tag{9}$$
The subassemblers’ pricing game can then be expressed as a game in which each subassembler $i$ selects a capacity level $0 \leq Q_i \leq Q_i(\alpha_i)$ and faces expected profit given by
\[
\pi_i(Q_i|Q_{j,j\neq i}, \alpha) = -(k_{i0} + k_i)Q_i^* + \left(\alpha_i - C^i - \sum_{j=1}^{m_i} k_{ij} \frac{F(Q_i)}{F(Q_i)} \right) E \left[\min\{Q_d^*(\alpha, \gamma), D\}\right]. \tag{10}
\]

Define now the isolated subassembler pricing problem to be that in which each subassembler $i$ can select the capacity for the entire system – i.e., subassembler $i$ selects capacity $0 \leq Q_i \leq Q_i(\alpha_i)$ to maximize
\[
\pi_i^S(Q_i|\alpha_i) = -(k_{i0} + k_i)Q_i + \left(\alpha_i - C^i - \sum_{j=1}^{m_i} k_{ij} \frac{F(Q_i)}{F(Q_i)} \right) E \left[\min\{Q_i, D\}\right] \\
\quad = -(k_{i0} + k_i)Q_i + \left(\alpha_i - C^i \right) \int_0^{Q_i} \bar{F}(x)dx - \sum_{j=1}^{m_i} k_{ij} \int_0^{Q_i} \bar{F}(x)dx.
\]

We next define
\[
l(Q) \overset{\text{def}}{=} 1 + \frac{f(Q)}{F(Q)} \int_0^Q \bar{F}(x)dx.
\]
The function $l(Q)$ is strictly increasing if and only if
\[
2 \frac{f(Q)}{F(Q)} + \frac{f'(Q)}{f(Q)} + \frac{\bar{F}(Q)}{\int_0^Q \bar{F}(x)dx} > 0. \tag{11}
\]
Condition (11) is satisfied, for example, by any distribution with an increasing failure rate (IFR), i.e., for which $f(Q)/\bar{F}(Q)$ is increasing in $Q$. Assume for the remainder of the paper that the distribution of demand satisfies (11). Observe that $l(0) = 1$ and $l(\cdot)$ is strictly increasing, which implies that $l(Q) > 1$ for all $Q > 0$.

Proposition 4 in Gerchak and Wang (2002) and Lemma 25 in Tomlin (2000) show that, under (11), $\pi_i^S(\cdot|\alpha_i)$ is concave and
\[
\frac{\partial \pi_i^S(Q_i|\alpha_i)}{\partial Q_i} = -(k_{i0} + k_i) + \left(\alpha_i - C^i \right) \frac{\bar{F}(Q_i)}{F(Q_i)} - \sum_{j=1}^{m_i} k_{ij} l(Q_i) = 0 \tag{12}
\]
has a unique interior solution $Q_i^S(\alpha_i)$. As a result, there exists a unique optimal capacity level $0 < Q_i^S(\alpha_i) < Q_i(\alpha_i)$ for subassembly $i$’s isolated pricing problem.

Define
\[
Q^{S*}(\alpha) = \min\{Q_1^S(\alpha_1), \cdots, Q_n^S(\alpha_n), Q_0^S(\alpha_0)\},
\]
and note that $0 < Q^{S*}(\alpha) < \min_{1 \leq i \leq n} Q_i(\alpha_i)$. We now show that the subassemblers’ pricing game, in its capacity form, has a Nash equilibrium.
Proposition 2 A set of subassembler capacity choices \{Q_1, \cdots, Q_n\} is a Nash equilibrium for the subassemblers’ pricing game with profit (10) if and only if \(Q_1 = \cdots = Q_n \leq Q^S_*(\alpha)\), and the equilibrium with \(Q_i = Q^S_*(\alpha)\) for \(i = 1, \ldots, n\) is uniquely Pareto-optimal for the subassemblers.

We conclude the subassemblers’ pricing game with the following Corollary that results from Proposition 2, (3), (6) and (9).

Corollary 1 Fix \(\alpha\). In the subassemblers’ pricing game, the subassemblers’ equilibrium price shares \(\gamma^S_*(\alpha)\) are given by

\[
\gamma^S_{ij}(\alpha) = \frac{c_{ij}}{\alpha_i} + \frac{k_{ij}}{\alpha_i F(Q^S_*(\alpha))},
\]

for all \(i = 1, \ldots, n\) and \(j = 1, \ldots, m_i\). Also, \(Q^*_d(\alpha, \gamma^S_*(\alpha)) = Q^S_*(\alpha)\).

Finally, note that \(Q^S_*(\alpha) > 0\) implies that \(F(Q^S_*(\alpha)) < 1\) which, in turn, implies that \(\alpha_i \gamma^S_{ij}(\alpha) > c_{ij} + k_{ij}\). At the same time, \(Q^S_*(\alpha) < \frac{\alpha_i}{\alpha_i - C_i}\) implies that \(F(Q^S_*(\alpha)) > \frac{K_i}{\alpha_i - C_i}\), so that \(\alpha_i \gamma_{i0} > c_{i0} + c_i + (k_{i0} + k_i)\alpha_i - C_i / K_i > c_{i0} + c_i + k_{i0} + k_i\), by (1). Thus, for any \(\alpha\) offered by the assembler, all suppliers and subassemblers make, at equilibrium, a positive margin for each unit sold.

3.3 Assembler’s Pricing Problem

In the assembler’s pricing problem, the assembler acts as the Stackelberg leader by selecting a vector of prices \(\alpha\) in anticipation of the subassemblers’ equilibrium price shares \(\gamma^S_*(\alpha)\) and the second stage capacity equilibrium \(Q^*_d(\alpha, \gamma^S_*(\alpha)) = Q^S_*(\alpha)\). Thus, the profit for the assembler is given by:

\[
\pi_0(\alpha) = -k_0 Q^S_*(\alpha) + (\alpha_0 - c_0) E[\min\{Q^S_*(\alpha), D\}]. \tag{13}
\]

We will show that it is optimal for the assembler to set \(\alpha\) so that \(Q^S_*(\alpha) = Q^S_1(\alpha_1) = \cdots = Q^S_n(\alpha_n)\). As a first step, we show that in each subassembler \(i\)’s isolated pricing problem the optimal capacity level is increasing in \(\alpha_i\).

Lemma 2 The optimal capacity level for subassembler \(i\)’s isolated pricing problem, \(Q^S_*(\alpha_i)\), is increasing in \(\alpha_i\). Furthermore, define

\[
\gamma^S_{ij}(\alpha_i) = \frac{c_{ij}}{\alpha_i} + \frac{k_{ij}}{\alpha_i F(Q^S_*(\alpha_i))}; \quad \gamma^S_{i0}(\alpha_i) = 1 - \sum_{j=1}^{m_i} \gamma^S_{ij}(\alpha_i).
\]
Then $\alpha_i \gamma^S_{ij}(\alpha_i)$, the payment to supplier $ij$ per unit sold, and $\alpha_i \gamma^S_{0i}(\alpha_i)$, are increasing in $\alpha_i$, i.e., all firms in module $i$ receive a higher payment for each unit sold as a result of an increase in the unit payment $\alpha_i$.

We can now prove the following result.

**Proposition 3** It is optimal for the assembler to choose the vector $\alpha$ so that:

(a) $Q^S_i(\alpha_i) \leq Q^0_i(\alpha_0)$ for all $i = 1, \ldots, n$,

(b) $Q^{S*}(\alpha) = Q^S_1(\alpha_1) = \cdots = Q^S_n(\alpha_n)$.

Proposition 3 allows us to reduce the assembler’s pricing problem to one in which he only selects one price, say $\alpha_1$, to maximize his profit. From (12) and part (b) of Proposition 3, it follows that

$$\alpha_i(\alpha_1) = C^i + \frac{k_{i0} + k_i + l(Q^S_1(\alpha_1)) \sum_{j=1}^{m_i} k_{ij}}{F(Q^S_1(\alpha_1))}, \ i > 1. \quad (14)$$

We now derive bounds on the optimal choice of $\alpha_1$. To that end, define $\alpha_0(\alpha_1) = 1 - \alpha_1 - \sum_{i=2}^{n} \alpha_i(\alpha_1)$ and note that $\alpha_0(\cdot)$ is decreasing. Recall that $\alpha_i^{min} = C^i + K^i$ and observe that $\alpha_i(\alpha_i^{min}) = C^i + K^i = \alpha_i^{min}$, so that $\alpha_0(\alpha_i^{min}) = 1 - \sum_{i=1}^{n} (C^i + K^i) > c_0 + k_0$. This implies that $0 = Q^S_1(\alpha_1^{min}) < Q^0_1(\alpha_0(\alpha_1^{min}))$. Define $\alpha_1^{max}$ to satisfy $Q^S_1(\alpha_1^{max}) = Q^0_1(\alpha_0(\alpha_1^{max}))$. Part (a) of Proposition 3 then implies that the assembler’s optimal choice of $\alpha_1$ will be between $\alpha_1^{min}$ and $\alpha_1^{max}$. In addition, observe that $\alpha_0(\alpha_1^{max}) > c_0 + k_0$. The assembler’s pricing problem then becomes

$$\max_{\alpha_1^{min} \leq \alpha_1 \leq \alpha_1^{max}} \pi_0(\alpha_1) = -k_0 Q^S_1(\alpha_1) + (\alpha_0(\alpha_1) - c_0) E[\min\{Q^S_1(\alpha_1), D\}]. \quad (15)$$

Unfortunately, it is hard to verify whether or not the profit function $\pi_0(\alpha_1)$ is unimodal in $\alpha_1$. However, it is a continuous function maximized over a closed interval. Thus, the function $\pi_0(\alpha_1)$ achieves its maximum at some $\alpha^*_1$, with $\alpha_1^{min} \leq \alpha^*_1 \leq \alpha_1^{max}$, and $\alpha_0(\alpha^*_1) > c_0 + k_0$.

The following Theorem summarizes the results for the multi-stage capacity and pricing game with modular assembly.

**Theorem 1** For a given set of unit capacity cost parameters $k_0, k_i, k_{ij}, i = 1, \ldots, n, j = 0, 1, \ldots, m_i$ and unit production/assembly cost parameters $c_0, c_i, c_{ij}, i = 1, \ldots, n, j = 0, 1, \ldots, m_i$, there exists a price $\alpha^*_1$ that achieves the maximum in (15) with $Q^S_1(\alpha^*_1)$ the unique solution of $\partial \pi^S_1(Q|\alpha^*_1)/\partial Q = 0$, so that the first stage price equilibrium is given by

$$\alpha^*_1 = C^i + \frac{k_{i0} + k_i + l(Q^S_1(\alpha^*_1)) \sum_{j=1}^{m_i} k_{ij}}{F(Q^S_1(\alpha^*_1))}, \ i = 1, \ldots, n,$$
\[ \gamma_{ij}^* = \frac{c_{ij}}{\alpha_i^*} + \frac{k_{ij}}{\alpha_i^* F(Q_1^S(\alpha_i^*))}, \quad i = 1, ..., n, \quad j = 1, ..., m_i. \]

This implies a unit net revenue of \( \alpha_{0i}^* = 1 - \sum_{i=1}^{n} \alpha_i^* \) for the assembler, and price shares \( \gamma_{i0}^* = 1 - \sum_{j=1}^{m_i} \gamma_{ij}^* \) for each subassembler \( i \). The second stage capacity equilibrium is given by:

\[ Q_d^*(\alpha^*, \gamma^*) = Q_1^S(\alpha_1^*). \]

Finally,

\[ \alpha_i^*\gamma_{ij}^* > c_{ij} + k_{ij}, \quad \alpha_i^*\gamma_{i0}^* > c_{i0} + c_i + k_{i0} + k_i \quad \text{and} \quad \alpha_0^* > c_0 + k_0. \]

That is, all firms in the system make, at equilibrium, a positive margin for each unit sold.

4 Design of the Modular Assembly System

The preceding section provides a (partial) characterization of the optimal assembler prices, the subsequent equilibrium subassembler price shares, and the resulting equilibrium capacities for a given module and cost structure. When deciding whether to shift from traditional to modular assembly, however, the assembler may have flexibility in structuring the modular assembly system. For example, the product architecture may allow for choice in terms of the relative balance among modules (e.g., the number of components in each module), or which supplier to choose as the subassembler for each module. In this section we use the results from the previous section to obtain insights that can guide these managerial design decisions.

We first consider the issue of balance among modules. As an example to illustrate this question, suppose the assembler could feasibly split the product into either of the two-module configurations shown in Figure 2.

![Figure 2: Balance](image-url)
Assume that in both configurations the same two suppliers are chosen as subassemblers for modules 1 and 2, respectively. Would the assembler prefer configuration A (unbalanced) or configuration B (balanced) – i.e., which one would yield higher optimal expected assembler profit? Proposition 4 below establishes a result that can help answer that question.

Define \( K_{SA} = \sum_{i=1}^{n} (k_i 0 + k_i) \) and \( C_{SA} = \sum_{i=1}^{n} (c_i 0 + c_i) \) to be the total unit capacity and production costs, respectively, for the subassemblers, and let \( K_S = \sum_{i=1}^{n} \sum_{j=1}^{m_i} k_{ij} \) and \( C_S = \sum_{i=1}^{n} \sum_{j=1}^{m_i} c_{ij} \) be the total costs for the suppliers.

**Proposition 4** For assembler prices \( \alpha \) satisfying Proposition 3, a modular assembly system is equivalent to a pure series system consisting of a single subassembler with unit costs \( K_{SA} \) and \( C_{SA} \) and a single supplier with unit costs \( K_S \) and \( C_S \), in that both systems yield the same system-wide equilibrium capacity and the same profit for the assembler. This holds, in particular, for the profit-maximizing assembler prices \( \alpha^* \) of Theorem 1.

Proposition 4 allows us to directly answer the balance question posed above in cases where changing components from one module to another does not change the total assembly-related costs at the subassemblers, nor does it change the costs incurred by the assembler – i.e., \( \sum_{i=1}^{n} k_i \), \( \sum_{i=1}^{n} c_i \), \( k_0 \) and \( c_0 \) are constant across the system configurations being considered. (Of course, even though the total costs are unchanged, shifting components from module \( i \) to module \( r \) may cause \( k_i \) and \( c_i \) to decrease while \( k_r \) and \( c_r \) increase.) This condition would hold if the assembly costs \( k_i \) and \( c_i \) for module \( i \) were determined by the specific components that make up the module – e.g., if physically attaching any particular component to a module costs the same (in terms of labor/machine capacity, as well as variable labor/material costs) regardless of which other components are part of that module.

In that event, Proposition 4 implies that module balance does not matter to the assembler – e.g., the assembler is indifferent between configurations A and B in Figure 2. Once a specific set of subassemblers is chosen, any assignment of components (and suppliers) to modules will result in a system configuration that can be reduced to the same series system. As a result, assembler profit will be the same under any such configuration.

There may be some cases in which the above assumption does not hold – i.e., the subassemblers’ total assembly-related costs are a function of how the modules are configured. One possible reason is economies of scale in assembly. For example, suppose that the costs \( k_i \) and \( c_i \) of assembling module \( i \) were given by \( k_i = \sum_{j=1}^{m_i} k^j_i + g_i(m_i) \) and \( c_i = \sum_{j=1}^{m_i} c^j_i + g_i(m_i) \), where \( k^j_i \) and \( c^j_i \) are costs specific to each component \( j \) in the module, and \( g_i(m_i) \) is a concave increasing function of the number of components in the module. In this case, the subassemblers’ total assembly-related costs would equal \( \sum_{i=1}^{n} k_i = \sum_{i=1}^{n} \sum_{j=1}^{m_i} k^j_i + \sum_{i=1}^{n} g_i(m_i) \) and
\[ \sum_{i=1}^{n} c_i = \sum_{i=1}^{n} \sum_{j=1}^{m_i} c_{ij}^i + \sum_{i=1}^{n} g_i(m_i). \] These quantities are minimized by setting \( m_i = 1 \) for all but one subassembler – i.e., by choosing the most unbalanced configuration possible. Proposition 4 would then imply that such a configuration is equivalent to a series system with lower \( K_{SA} \) and \( C_{SA} \) than any other way of assigning components to those modules. Since lower subassembler costs yield higher assembler profit, the assembler would prefer this unbalanced configuration.

In other cases, the relationship between component allocations to modules and the subassemblers’ total assembly-related costs may be more complex, so that unambiguous comparisons between the resulting \( K_{SA} \) and \( C_{SA} \) costs may not be possible. (For example, there may be economies of scale in assembly, combined with diseconomies of scale in managing supplier relationships.) In such cases, Proposition 4 can still help an assembler choose between alternative ways of allocating components to modules. The assembler simply needs to reduce each configuration to its equivalent series system and solve that simplified problem, then choose the configuration that yields higher expected profit.

**Remark 1** The result in Proposition 4 is also valid for a traditional assembly system in that, under an assembler’s payment \( \alpha \), the resulting equilibrium capacity and expected assembler profit for the general system are equal to those achieved in the corresponding pure series system, by requiring, in this case, that the payment \( \alpha \) satisfy a version of Lemma 1, with \( m_i = n \), \( \alpha_i = 1 \) and \( \gamma_{ij} = \alpha_j \). The proof for the case of a traditional assembly system closely resembles that of Proposition 4. This result for the single-tier traditional system was observed by Gerchak and Wang (2002).

Proposition 4 addresses the question of allocating components across modules given a fixed number of modules, each assembled by a specific subassembler. We now turn to the question of choosing a subassembler.

Consider any particular allocation of components into modules. Assume that each of the suppliers producing components for module \( i \) is equally capable of assembling the components to make the module at the same (capacity and unit assembly) cost. However, the suppliers may differ in terms of their component capacity and unit production costs \( k_{ij} \) and \( c_{ij} \), respectively. Which supplier should the assembler choose to take on the subassembler role? Proposition 5 below helps answer this question.

**Proposition 5** A shift in unit component production cost \( c \) from supplier to subassembler has no effect on the assembler’s optimal profit. A shift in unit component capacity cost \( k \) from supplier to subassembler causes the assembler’s optimal profit to increase.
Proposition 5 implies that the assembler can maximize his profit by choosing as subassembler the supplier with the largest component capacity cost $k_{ij}$ – thus shifting the largest possible amount of capacity cost from supplier to subassembler. (Note that supplier production costs $c_{ij}$ do not matter.) If some suppliers are incapable of performing assembly, the assembler merely needs to select the largest $k_{ij}$ from among those who are capable of assembly.

To see the intuition behind this result, consider the impact of $c$ and $k$ on the assembler. The unit costs $c$ do not represent risk to suppliers or subassemblers, since they are only incurred on items that are sold. The assembler only needs to cover those costs, plus offer a sufficiently large margin to induce suppliers and subassemblers to install the desired capacity in the face of the risky upfront costs $k$. Since the assembler can influence subassemblers directly through payment of $\alpha$, but can only influence suppliers indirectly (through the subassemblers, who pass on only a portion of $\alpha$), it is easier to overcome subassembler capacity costs than supplier costs. As a result, for each module $i$, the assembler prefers to make $\sum_{j=1}^{m_i} k_{ij}$ as small as possible – which is achieved by selecting the supplier with the highest capacity cost to be the subassembler.

Proposition 5 has two interesting implications. First, suppose the subassembler has the capability to eliminate one or more of the suppliers and produce those components herself (at the same cost). Such a move would correspond to shifting some of the $k_{ij}$ and $c_{ij}$ costs from the supplier tier to the subassembler tier, which would increase the assembler’s expected profits. Second, Proposition 5 suggests that the assembler’s optimal expected profit is higher when there are more modules. This is because a larger number of modules allows the component capacity cost $k_{ij}$ of one more supplier to be moved up to the subassembler tier. This observation, when taken to the limit, has interesting implications regarding the attractiveness of modular assembly relative to traditional assembly. This limiting case is analogous to traditional assembly, and this logic would imply that modular assembly always leads to lower optimal expected costs for the assembler. If we assume now that total assembly-related costs $\sum_{i=0}^{n} k_i$ and $\sum_{i=0}^{n} c_i$ are constant in the system, this conclusion is in fact correct – and it is made rigorous in the following two results. (When total assembly-related costs are not constant, we can compare the performances of traditional assembly and modular assembly by reducing each configuration to its equivalent series system, in view of Proposition 4, and solving and comparing those simplified problems.)

The next result provides additional insight regarding the impact on the assembler’s profit of shifting to modular assembly. Proposition 6 says that the more assembly the assembler
outsources, as measured by assembly capacity costs, the more the assembler’s profit will deteriorate.

**Proposition 6** A shift in unit assembly cost $c$ from assembler to subassembler has no effect on the assembler’s optimal profit. A shift in unit assembly capacity cost $k$ from assembler to subassembler reduces the assembler’s optimal profit.

We conclude by showing that modular assembly leads to inferior performance relative to traditional assembly.

**Theorem 2** If total assembly related costs in the system are held constant, then a modular assembly system leads to lower optimal expected profit for the assembler than a traditional system in which the assembler performs all the assembly operations. Furthermore, the resulting capacity in the modular assembly system is lower than the capacity that arises in the traditional system, and the aggregate expected profit for the entire assembly system is lower under modular assembly than under traditional assembly.

The result in Theorem 2 continues to hold even if there is no cost passed to the subassemblers – i.e., if the assembler pays for all the assembly operations outsourced to the subassembler.

5 **Profitable Modular Assembly Through Cost Reduction**

At first glance the results at the end of the previous section seem to imply that companies would be foolish to pursue modular assembly, since doing so would decrease their expected profit. While those results provide important cautions regarding the potential downside of modular assembly, they do not tell the full story. Assembler profit decreases if the subassemblers incur the same costs as the assembler to perform the outsourced work. However, in practice the subassemblers may face lower assembly costs than the assembler (due to lower hourly wages, greater efficiency, etc.). As a result, companies typically face a trade-off when considering modular assembly: outsourcing assembly to suppliers may drive some cost out of the system which would increase profitability, but doing so also results in a loss of assembler control which would lower assembler profit. The relative impacts of these two factors will determine whether a shift to modular assembly in any given situation is a profitable choice.
(i.e., yields a higher expected profit for the assembler than traditional assembly). In this section we explore this trade-off and obtain insights about the factors that determine when modular assembly is a good choice for the assembler.

In light of Proposition 4, without loss of generality we focus on a modular assembly system with a series structure. The single supplier is indexed \{11\}, the single subassembler is indexed \{10\}, and the assembler is indexed \{0\}. In the analogous traditional assembly system, the subassembler becomes a second supplier, which we still index \{10\}. The two systems are depicted in Figure 3.

![Assembly Systems Diagram](image)

**Figure 3: Assembly Systems**

Player \{10\} has the same component costs \(c_{10}\) and \(k_{10}\) whether it is a supplier or a subassembler – similarly, supplier \{11\} has the same costs \(c_{11}\) and \(k_{11}\) in both systems.

In the previous section the assembly costs incurred by the assembler (\(c_0\) and \(k_0\)) and the subassembler (\(c_1\) and \(k_1\)) were taken as given. Here, in order to make comparisons between traditional and modular systems when the subassembler has a cost advantage, we need to reinterpret and expand our notation somewhat. First, for a traditional assembly system let \(c_0\) and \(k_0\) be the unit assembly and assembly capacity costs, respectively, incurred by the assembler. Now consider a shift to a modular assembly system. Such a shift results in the assembler outsourcing some of the assembly costs – call them \(c_\Delta\) and \(k_\Delta\) – leaving the assembler with costs \(c_0 - c_\Delta\) and \(k_0 - k_\Delta\). (I.e., \(c_0 - c_\Delta\) and \(k_0 - k_\Delta\) play the role of \(c_0\) and \(k_0\) in preceding sections.) As before, the subassembler’s (player \{10\}’s) assembly-related costs are given by \(c_1\) and \(k_1\). However, to represent the fact that the subassembler may have an assembly cost advantage over the assembler, these costs satisfy \(c_1 \leq c_\Delta\) and \(k_1 \leq k_\Delta\).

Throughout this section we will be interested in how much of an assembly-cost reduction the subassembler would need to provide in order to make modular assembly profitable. Define \(k^*(k_\Delta) \in [0, k_\Delta]\) to be the break-even subassembler assembly-capacity cost – i.e., if the assembler outsources assembly with (assembler) capacity costs of \(k_\Delta\) to the subassembler, then if it costs the subassembler \(k_1 = k^*(k_\Delta)\) to install the necessary assembly capacity, the
assembler’s profits would be the same under modular assembly as under traditional assembly.

In some cases \( k^*(k_\Delta) \) may not exist – i.e., even if the subassembler were able to install assembly capacity for free, modular assembly still would not be profitable. When \( k^*(k_\Delta) \) exists it will sometimes be convenient to refer to \( k_\Delta - k^*(k_\Delta) \), which represents the amount of cost reduction the subassembler must achieve to reach break-even. If the subassembler can achieve a cost reduction greater than \( k_\Delta - k^*(k_\Delta) \), then modular assembly is more profitable than traditional assembly. We define \( c^*(c_\Delta) \) and \( c_\Delta - c^*(c_\Delta) \) analogously for unit assembly costs.

We first consider an assembler that is deciding how much of the assembly task to outsource to a subassembler – where higher levels of outsourcing correspond to larger values of \( c_\Delta \) or \( k_\Delta \). Would an assembler outsourcing a larger amount of assembly require more or less assembly cost savings to make modular assembly profitable? In other words, how do \( k^*(k_\Delta) \) and \( k_\Delta - k^*(k_\Delta) \) \([c^*(c_\Delta) \text{ and } c_\Delta - c^*(c_\Delta)]\) behave as \( k_\Delta \) increases? The answer is not immediately clear, since passing more assembly capacity costs \( k_\Delta \) to a subassembler has two impacts on assembler profit. On one hand, Proposition 6 says that such a shift (ignoring for now the cost reduction the subassembler can achieve) causes assembler profit to decrease. On the other hand, passing more cost to the subassembler means that the subassembler’s cost advantage applies to a larger portion of total assembly costs – i.e., more total cost is driven out of the system. The following result shows that the latter factor always dominates – assemblers who outsource a larger amount of assembly-related costs find it easier to achieve break even, in the sense that their subassemblers need to deliver less of a relative assembly cost reduction.

**Proposition 7** Assume that the costs \( c_{10}, k_{10}, c_{11}, k_{11}, c_0 \) and \( k_0 \) are fixed. Then,

(a) One of the following occurs:

(i) Traditional assembly is more profitable for all \( 0 \leq k_\Delta \leq k_0 \), even if the subassembler can install the required assembly capacity for free (i.e., \( k^*(k_\Delta) \) does not exist for all \( 0 \leq k_\Delta \leq k_0 \)); or

(ii) There exists a threshold outsourcing level \( \hat{k}_\Delta \) such that traditional assembly is more profitable for all \( k_\Delta < \hat{k}_\Delta \), but for \( \hat{k}_\Delta \leq k_\Delta \leq k_0 \) there exists a break-even subassembler capacity cost \( k^*(k_\Delta) \).

(The same result holds for the unit assembly cost \( c \).)

(b) In case (a)(ii), as \( k_\Delta \) increases from \( \hat{k}_\Delta \) to \( k_0 \), the absolute amount of cost reduction \( k_\Delta - k^*(k_\Delta) \) required to achieve break even increases, and the relative cost reduction \( (k_\Delta - k^*(k_\Delta))/k_\Delta \) decreases.
In case (a)(ii), as $c_\Delta$ increases from $\hat{c}_\Delta$ to $c_0$, the absolute amount of cost reduction $c_\Delta - c^*(c_\Delta)$ required to achieve break even remains constant, and the relative cost reduction $(c_\Delta - c^*(c_\Delta))/c_\Delta$ decreases.

While it was convenient to state the preceding result using generic cost parameters, a more detailed discussion of the factors contributing to those costs (and differences between them) may facilitate interpretation of the result.

As mentioned above, one way a cost advantage may arise is if the subassembler’s hourly wage, say $w_1$, is lower than that paid by the assembler, $w_0$. If labor contracts are flexible (so that labor is a variable cost), then outsourcing a portion of assembly requiring $l_0$ hours of labor per unit yields cost parameters $c_\Delta = w_0l_0$ and $c_1 = w_1l_0 < c_\Delta$. (This assumes, for brevity of exposition, that variable assembly costs consist entirely of labor costs.) If instead labor contracts are long-term (so that labor is a fixed cost), the same would hold instead for unit capacity costs – i.e., $k_\Delta = w_0l_0$ and $k_1 = w_1l_0 < k_\Delta$. Consider the latter case. Part (a) of Proposition 7 says that if very little assembly is outsourced ($k_\Delta$ is small), then it is possible (case i) that even free labor ($w_1 = 0$) would not be enough to make modular assembly profitable. Note that what matters here is the size of $k_\Delta$ – the specific combination of $w_0$ and $l_0$ associated with that $k_\Delta$ does not matter. For example, if a large number of labor hours $l_0$ is outsourced, but the wage $w_0$ paid by the assembler is very small, we can interpret case (i) to mean that there is not enough room below $w_0$ for the subassembler to drop wages to achieve break-even. If instead $w_0$ is larger but $l_0$ is very small, we can interpret case (i) to say that even if the subassembler pays a much lower wage, that wage advantage gets applied to a very small number of hours, so the total savings is not sufficient to achieve break-even.

The other possibility (case ii) is that as long as at least some threshold amount of work is outsourced ($k_\Delta$ is large), then there exists an hourly wage, defined by $k^*(w_0l_0)/l_0$ such that modular assembly is profitable if the subassembler’s hourly wage satisfies $w_1 \leq k^*(w_0l_0)/l_0$. This relationship can be rewritten as

$$1 - w_1/w_0 \geq 1 - k^*(w_0l_0)/w_0l_0. \quad (16)$$

Part (b) of Proposition 7 says that the right-hand side of (16) is decreasing in $k_\Delta = w_0l_0$. So as more work is outsourced, (16) will hold for higher and higher subassembler wages $w_1$ (relative to assembler wages $w_0$) – i.e., less and less of a relative cost advantage.

A similar interpretation applies if instead both the subassembler and the assembler pay the same wages $w_0$, but the subassembler can perform assembly more efficiently – e.g., using fewer labor hours $l_1 < l_0$. (This may happen if the subassembler has greater experience
and expertise in handling their own component and similar components that are paired in a module – e.g., an engine and transmission.) The interpretation applies equally well to the case of variable costs $c_\Delta$, and can also be extended to situations in which some non-labor costs are also outsourced as part of $c_\Delta$ and $k_\Delta$.

We next explore whether modular assembly is more or less likely to be profitable in a system with high assembly costs (vs. a lower-assembly-cost system). To address this question, we examine the impact of changing $c_0$ or $k_0$ on the subassembler’s break-even cost level. To avoid introducing the effects described in Proposition 7, we keep $c_\Delta$ and $k_\Delta$ fixed as we change $c_0$ or $k_0$ – i.e., the absolute amount of assembly cost being outsourced is kept constant (although the relative amount will be different for different values of $c_0$ and $k_0$). Since now it is changes in $c_0$ or $k_0$ that will affect the subassembler’s break-even cost, we will use the modified notation $c^*(c_0)$ and $k^*(k_0)$ to refer to those quantities.

Similar to before, assembly costs are made up of several components including labor hours and hourly wages. For example, system A could have lower assembly capacity costs than system B if both produce a product requiring the same number of assembly hours, but system A pays lower wages – i.e., $k^A_0 = w^A_0 l_0 < w^B_0 l_0 = k^B_0$ – or if wages are the same but assembly hours differ – i.e., $k^A_0 = w_0 l^A_0 < w_0 l^B_0 = k^B_0$. As in Propositon 7, it is the total assembly capacity cost $k_0$ (unit assembly cost $c_0$) that affects the profitability of modular assembly, not the specific factors that contribute to that cost. As a result, our analysis focuses on changes in $k_0$ and $c_0$ – more detailed interpretations in terms of underlying factors can then be made similar to above.

Evaluating changes in $c_0$ and $k_0$ turns out to be more complex than for $c_\Delta$ and $k_\Delta$. As a result, we explore this issue through a numerical study. The values of the parameters used in the numerical study were not selected to represent any specific realistic setting. Rather, these values were chosen to compare expected profit achieved by the assembler in the traditional and modular assembly systems across a range of parameter values.

In all scenarios studied, finished product demand has a Normal distribution with a mean of 20 and a standard deviation of 5. Component cost parameters are:

- Supplier costs $c_{11} = k_{11} = 0.05$
- Supplier/subassembler costs $c_{10} = 0.05; k_{10} = 0.05, 0.10, 0.25$

The study examines a large number of specific problems generated by using the three $k_{10}$ values above, combined with a wide range of values for $c_0$ and $k_0$. For each such problem, we compute the subassembler’s break-even cost.

The first part of the study focuses on the impact of assembly capacity costs. We fix
$c_\Delta = 0$ and consider $c_0$ values of 0.05, 0.10 and 0.25. These three values, combined with the three values of $k_{10}$ identified above, yield 9 scenarios. For each scenario, the value of $k_\Delta$ is fixed at 0.05, while $k_0$ varies over the range 0.05, 0.075, ..., 0.25, yielding a total of 81 individual problems. The second part of the study conducts similar trials, this time focusing on the unit assembly cost $c_0$. In this part $k_\Delta = 0$ and $k_0$ can take on values 0.05, 0.10 or 0.25, while $c_\Delta$ is fixed at 0.05 and $c_0$ varies over the range 0.05, 0.075, ..., 0.25.

In part 1, the break-even cost level $k^*(k_0)$ could be found in 31 out of the 81 problems considered. In other words, in 50 out of 81 problems, assembler profit was lower under modular assembly than under traditional assembly, even if all of the $k_\Delta = 0.05$ cost were driven out of the system – i.e., if the subassembler could do the assembly for free. Similar results held for part 2 of the study – $c^*(c_0)$ could be found in 32 out of the 81 problems considered. In all cases, the inability to achieve break even corresponded to small values of $k_0$ or $c_0$. Once $k_0$ or $c_0$ increased to some threshold level such that break even could be achieved, it could always be achieved for larger values of $k_0$ or $c_0$ as well.

Restricting attention to those problems where $k^*(k_0)$ or $c^*(c_0)$ could be found, the cost reductions $k_\Delta - k^*(k_0)$ and $c_\Delta - c^*(c_0)$ were found to be decreasing in $k_0$ and $c_0$, respectively. In other words, assemblers with high initial assembly costs found it easier to achieve break even, in the sense that their subassemblers needed to deliver less of an assembly cost advantage.

Finally, we performed one more pair of numerical studies – this time exploring the impact of demand variability on the profitability of modular assembly. For both studies we used the baseline cost parameters given above (with $k_{10} = 0.05$) and set $c_0 = k_0 = 0.25$. In the first study we set $c_\Delta = 0$ and $k_\Delta = 0.1$, and computed the break-even subassembler assembly capacity cost (now denoted $k^*(\sigma)$) for each value of demand standard deviation, which took on values $\sigma = 1, 2, ..., 6$. In the second study we set $c_\Delta = 0.1$ and $k_\Delta = 0$, and computed the break-even subassembler unit assembly cost (now denoted $c^*(\sigma)$) for each value of $\sigma$.

In both studies we found that the break-even cost was decreasing in $\sigma$. As demand became more variable, the subassembler needed to provide a greater cost advantage in order to make modular assembly profitable. This makes intuitive sense. Demand uncertainty at the time capacity decisions are made (and the resulting risk faced by suppliers), combined with only an indirect ability to influence supplier capacity through module prices, is what causes the assembler to offer higher price incentives under modular assembly than he would under traditional assembly. The numerical study suggests that greater demand variability exacerbates this disadvantage of modular assembly.
6 Conclusions

In this paper, we developed a multi-tier assembly model representing the type of modular assembly environment that is becoming increasingly common in industry. For such a system we established the existence and characterized properties of equilibrium pricing and capacity decisions by suppliers, subassemblers and the final assembler. We showed that, in equilibrium, the assembler and the subassemblers always set prices so that all suppliers choose the same capacity level, which determines the system capacity.

Building on these fundamental results, we examined higher-level questions about how the assembler would like to design the modular assembly system. We found that, in the absence of economies of scale in assembly, the assembler is indifferent between a balanced and an unbalanced system; with economies of scale, the assembler would prefer an unbalanced system. For a given modular structure, the assembler should choose the supplier with the highest component capacity cost in each module to be the subassembler for that module. We also showed that, if the subassemblers do not have lower assembly-related costs than the assembler, then modular assembly results in strictly lower expected profit for the assembler as compared to traditional assembly. This profit gap increases as the assembler outsources more assembly work to the subassemblers. In addition, modular assembly results in lower system capacity and lower expected system-wide profit.

When subassemblers have lower assembly-related costs than the assembler, these cost savings have the potential to counter the profit loss the assembler experiences from outsourcing some of the assembly. We showed that when cost savings are present, modular assembly is more likely to be profitable when a large portion of total assembly cost is outsourced by the assembler. The results of numerical studies suggested that profitable modular assembly is also more likely when assembly costs represent a large portion of total system costs, and is less likely when demand variance is high.

By choosing a relatively simple model that focuses on certain aspects of modular assembly, we have been able to obtain strong results regarding the impacts of those aspects. This raises the question, however, of how our results might extend to settings that exhibit additional complexities not contained in our model. Although rigorous analysis of such settings is beyond the scope of this paper, we provide a brief discussion and some conjectures.

In some situations contracts between the players may last significantly longer than the lead time to add capacity. In such cases, multiple opportunities to add capacity would exist over the duration of the contract (which might now be the natural time horizon to model). However, it seems that similar behavior would occur in this setting as well, because
the same forces are at work. The assembler and subassemblers would set prices to achieve balance among critical fractiles (up front) and thus capacity choices (every time capacity is added), but the assembler would still only be able to influence suppliers indirectly through prices offered to the subassemblers. This would lead to lower capacities and (assembler and supply chain) profits than a traditional system. Similarly, the same basic logic used to establish the other results (module balance does not matter, high assembly-capacity-cost suppliers should be chosen as subassemblers) appear to apply here as well. The size of the performance gap between modular and traditional assembly may be different, however. With multiple capacity expansion opportunities (and presumably more accurate demand information as time passes), players would probably install less capacity up front, thus reducing the risk of excess capacity. This may narrow the performance gap (similar to our numerical result that low demand variability enhances profitability of modular assembly), but would not eliminate it. In addition, if initial contracts spanning multiple periods could not be completely specified, the firms would face the hold-up and renegotiation problems analyzed in the economics literature.

Finally, there may be other factors that affect the performance gap between traditional and modular assembly. For example, in reality supplier/subassembler control over capacity may lie somewhere between voluntary and forced compliance (where the assembler can essentially dictate all actions and capture all profits). While the performance gap between modular and traditional assembly would still exist in such a setting, it may be smaller. On the other hand, longer-term labor dynamics might make modular assembly less attractive. As subassemblers grow in size and importance with their new responsibilities, they may become more attractive candidates for unionization and/or labor disruptions (Slaughter 1999). This could reduce potential cost advantages at the subassemblers and thus make modular assembly less attractive. All of these issues represent potential directions for future research.

Acknowledgements. The authors thank Paul Zipkin, Brian Tomlin and participants in seminars at Duke University, The University of Chicago, The University of Texas, and the MSOM 2003 Conference at USC for feedback on earlier versions of this work. We also thank the associate editor and referees for their helpful comments. Finally, we thank Tracy Lewis for suggestions regarding literature on the boundaries of the firm.

References


Appendix: Proofs

Proof of Proposition 1. Consider any Nash equilibrium $Q^*$. For any two players $h_1$ and $h_2$, $Q_{h_1}^* \leq Q_{h_1}(Q^*) \leq Q_{h_2}^* \leq Q_{h_2}(Q^*) \leq Q_{h_1}^*$, so that $Q_{h_1}^* = Q_{h_2}^*$. Also, $Q_h^* \leq Q_h^I$ for all $h$, so that $Q_h^* \leq Q_d^*$. Wang and Gerchak (2003) show that any $Q$ such that $Q_{ij} = Q_i = Q_0 \leq Q_0^*(\alpha, \gamma)$ is a Nash equilibrium and that $Q_0^*$ is uniquely Pareto-optimal. ■

Proof of Proposition 2. Given any capacity choices by the other subassemblers, if subassembler $i$ chose $Q_i > \min \{\min_{x \neq i} Q_x, Q_0^I(\alpha)\}$, then she could increase her expected profit by reducing her capacity to $Q_i = \min \{\min_{x \neq i} Q_x, Q_0^I(\alpha)\}$, thus increasing $\gamma_i$ (see (8)). Also, for $Q_i < \min \{\min_{x \neq i} Q_x, Q_0^I(\alpha)\}$ subassembler $i$’s expected profit is the same as in its isolated pricing problem, so if $Q_i^S(\alpha_i) < \min \{\min_{x \neq i} Q_x, Q_0^I(\alpha)\}$ it is optimal to choose $Q_i = Q_i^S(\alpha_i)$. As a result, subassembler $i$’s best-response capacity in this game is $\min \{\min_{x \neq i} Q_x, Q_0^I(\alpha), Q_i^S(\alpha_i)\}$. It is then easy to verify that any $\{Q_1, \ldots, Q_n\}$ such that $Q_1 = \cdots = Q_n \leq Q^S(\alpha)$ is a Nash equilibrium for the subassemblers’ pricing game. Similar to the proof of Proposition 1, consider now $Q^*$ a Nash equilibrium. For any two players $h_1$ and $h_2$, $Q_{h_1}^* = \min \{\min_{x \neq h_1} Q_x^*, Q_0^I(\alpha), Q_i^S(\alpha_i)\} \leq Q_{h_2}^* = \min \{\min_{x \neq h_2} Q_x^*, Q_0^I(\alpha), Q_i^S(\alpha_i)\} \leq Q_{h_1}^*$, so that $Q_{h_1}^* = Q_{h_2}^* \leq Q^S(\alpha)$.

Since each subassembler $i$’s expected profit is increasing on $Q_i \leq Q^S(\alpha)$, the equilibrium with $Q_i = Q^S(\alpha)$ for $i = 1, \ldots, n$ is uniquely Pareto-optimal. ■

Proof of Lemma 2. Observe from (12) that $\frac{\partial^2 \pi_i^S(Q)}{\partial Q^2} = \overline{F}(Q_i) > 0$. Also, from (12) and the fact that $I(\cdot)$ is increasing, we have that

$$\frac{\partial^2 \pi_i^S(Q)}{\partial Q^2} = -(\alpha_i - C^i) f(Q_i) - \sum_{j=1}^{m_i} k_{ij} I'(Q_i) \leq -(\alpha_i - C^i) f(Q_i).$$ (17)

Since $Q_i^S(\alpha_i)$ is the solution of (12), we can compute $\frac{\partial Q_i^S}{\partial \alpha_i} = -\frac{\partial^2 \pi_i^S/\partial Q^2}{\partial^2 \pi_i^S/\partial Q^2}$ by the Implicit Function Theorem, and conclude, from (17), that

$$0 \leq \frac{\partial Q_i^S}{\partial \alpha_i} \leq \frac{\overline{F}(Q_i^S)}{(\alpha_i - C^i) f(Q_i^S)}. $$

We thus conclude that $Q_i^S(\alpha_i)$ is increasing in $\alpha_i$, and this increase can only be achieved by subassembler $i$ passing on to her suppliers some of the price increase she receives from the assembler – i.e., $\alpha_i \gamma_{ij}^S(\alpha_i)$ is increasing in $\alpha_i$. It is also possible to identify an upper bound on how fast $\alpha_i \gamma_{ij}^S(\alpha_i) = c_{ij} + k_{ij}/\overline{F}(Q_i^S(\alpha_i))$ increases in $\alpha_i$, for $j = 1, \ldots, m_i$, that is

$$0 \leq \frac{\partial (\alpha_i \gamma_{ij}^S)}{\partial \alpha_i} = \frac{k_{ij} f(Q_i^S) (\partial Q_i^S/\partial \alpha_i)}{\overline{F}(Q_i^S)^2} \leq \frac{k_{ij}}{\overline{F}(Q_i^S)(\alpha_i - C^i)}. $$
This in turn implies that
\[
\frac{\partial (\alpha_i \gamma_i^S)}{\partial \alpha_i} = 1 - \sum_{j=1}^{m_i} \frac{\partial (\alpha_i \gamma_{ij}^S)}{\partial \alpha_i} \geq 1 - \frac{\sum_{j=1}^{m_i} k_{ij}}{F(Q_i^S(\alpha_i - C_i))} = 1 - \frac{\alpha_i \sum_{j=1}^{m_i} \gamma_{ij}^S - \sum_{j=1}^{m_i} c_{ij}}{\alpha_i - C_i} = \frac{\alpha_i \gamma_{i0}^S - c_{i0} - c_i}{\alpha_i - C_i}.
\]

The last expression can be shown to be positive as follows: (3), (4) and (5) imply that for \( j = 1, \ldots, m_i, \)
\[
\frac{c_{ij} + k_{ij}}{\alpha_i} \leq \gamma_{ij}^S \leq \frac{k_{ij}}{K_i} + \frac{c_{ij}}{\alpha_i} - \frac{k_{ij} C_i}{\alpha_i K_i}.
\]

Taking the second inequality in (18), summing both sides over \( j = 1, \ldots, m_i \) and noting that
\[
\alpha_i - \alpha_i \gamma_{i0}^S = \alpha_i \sum_{j=1}^{m_i} \gamma_{ij}^S,
\]
implies \( \alpha_i \gamma_{i0}^S \leq c_{i0} + c_i + \frac{k_{i0} + k_i}{K_i} (\alpha_i - C_i). \) We then conclude that
\[
\frac{\partial (\alpha_i \gamma_{i0}^S)}{\partial \alpha_i} \geq \frac{\alpha_i c_{i0} - c_i}{\alpha_i - C_i} \geq \frac{k_{i0} + k_i}{K_i} > 0.
\]

**Proof of Proposition 3.** (a) Suppose the assembler sets \( \alpha \) such that \( Q_j^S(\alpha_j) > Q_0^I(\alpha_0) \geq Q_i^S(\alpha_i) \) for any two subassemblers \( i \) and \( j. \) By Lemma 2, \( Q_j^S(\alpha_j) \) is increasing in \( \alpha_j. \) Also, \( Q_0^I(\alpha_0) \) is increasing in \( \alpha_0, \) i.e., it is decreasing in \( \alpha_j. \) Thus, by slightly decreasing \( \alpha_j \) to \( \hat{\alpha}_j \) so that \( Q_j^S(\hat{\alpha}_j) > Q_0^I(\hat{\alpha}_0) > Q_i^S(\alpha_i), \) the system capacity does not change and \( \hat{\alpha}_0 > \alpha_0, \) where \( \hat{\alpha}_0 \) reflects the change in \( \alpha_j. \) This implies an increase in the assembler’s profit (13).

Suppose now that the assembler sets \( \alpha \) such that \( Q_j^S(\alpha_j) > Q_0^I(\alpha_0) \) for all \( j = 1, \ldots, m_i. \) Then, \( Q^S(\alpha) = Q_0^I(\alpha_0) \) so that \( \pi_0(\alpha) = \min_{Q_0} \{-k_0 Q_0 + (\alpha_0 - c_0) E[\min\{Q_0, D\}]} \), which is increasing in \( \alpha_0. \) Therefore, by slightly decreasing \( \alpha_j, \) the assembler’s profit increases.

(b) Suppose that the assembler selects \( \alpha \) so that \( Q_j^S(\alpha_i) < Q_j^S(\alpha_j) \leq Q_0^I(\alpha_0) \) for any two subassemblers \( i \) and \( j. \) Then, a slight decrease in \( \alpha_j \) to \( \hat{\alpha}_j \) implies \( Q_j^S(\alpha_i) \leq Q_j^S(\hat{\alpha}_j) < Q_0^I(\hat{\alpha}), \) so that \( Q^S(\alpha) = Q^S(\hat{\alpha}) \) and \( \hat{\alpha}_0 > \alpha_0, \) where \( \hat{\alpha} \) and \( \hat{\alpha}_0 \) reflect the change of \( \alpha_j \) to \( \hat{\alpha}_j. \) This implies an increase in the assembler’s profit (13). It is then optimal for the assembler to choose \( \alpha \) so that \( Q^S(\alpha) = Q_1^S(\alpha_1) = \cdots = Q_n^S(\alpha_n). \)

**Proof of Proposition 4.** Let \( \alpha \) be a vector of assembler’s prices for the general system satisfying Proposition 3 and let \( Q^*_d(\alpha, \gamma^S(\alpha)) \) be the resulting equilibrium system capacity. Then, Proposition 3 implies that for each \( i = 1, \ldots, n, \) \( Q^*_d(\alpha, \gamma^S(\alpha)) \) satisfies (12), i.e.,
\[
-(k_{i0} + k_i) + (\alpha_i - C_i) F(Q^*_d(\alpha, \gamma^S(\alpha))) - \sum_{j=1}^{m_i} k_{ij} l(Q^*_d(\alpha, \gamma^S(\alpha))) = 0.
\]

Summing these expressions over \( i \) yields
\[
-K_{SA} + \left( \sum_{i=1}^{n} \alpha_i - C_S - C_{SA} \right) F(Q^*_d(\alpha, \gamma^S(\alpha))) - K_{Sl}(Q^*_d(\alpha, \gamma^S(\alpha))) = 0.
\]
Note that (19) is just the first-order optimality condition for the series system if the assembler chooses a price \( \tilde{\alpha} = \sum_{i=1}^{n} \alpha_i \), so that choice of price yields an equilibrium capacity in the series system of \( \tilde{Q}_d(\tilde{\alpha}, \gamma^{S^*}(\tilde{\alpha})) = Q_d^*(\alpha, \gamma^{S^*}(\alpha)) \). This choice allows the assembler to achieve the same system capacity and the same profit margin as in the general system, so he will earn the same expected profit.

Conversely, consider now \( \tilde{\alpha} \), an assembler’s price in the series system, satisfying Proposition 3 (part (a)). The equilibrium capacity for the series system is then given by \( \tilde{Q}_d(\tilde{\alpha}, \gamma^{S^*}(\tilde{\alpha})) \), and by Proposition 3(a), \( \tilde{Q}_d(\tilde{\alpha}, \gamma^{S^*}(\tilde{\alpha})) \leq Q_0^I(\tilde{\alpha}_0) \), where \( \tilde{\alpha}_0 = 1 - \tilde{\alpha} \). The capacity \( \tilde{Q}_d(\tilde{\alpha}, \gamma^{S^*}(\tilde{\alpha})) \) satisfies

\[
-K_{SA} + (\tilde{\alpha} - C_S - C_{SA}) \tilde{F}(Q_d^*(\tilde{\alpha}, \gamma^{S^*}(\tilde{\alpha}))) - K_S l(Q_d^*(\tilde{\alpha}, \gamma^{S^*}(\tilde{\alpha}))) = 0. \tag{20}
\]

Following (14), define now

\[
\alpha_i = C^i + \frac{k_{i0} + k_i + l(\tilde{Q}_d^*(\tilde{\alpha}, \gamma^{S^*}(\tilde{\alpha}))) \sum_{j=1}^{m_i} k_{ij}}{\tilde{F}(Q_d^*(\tilde{\alpha}, \gamma^{S^*}(\tilde{\alpha})))}.
\]

Then, by its definition, \( Q_i^S(\alpha_i) = \tilde{Q}_d^*(\tilde{\alpha}, \gamma^{S^*}(\tilde{\alpha})) \), for \( i = 1, ..., n \), and

\[
\sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} C^i + \frac{1}{\tilde{F}(Q_d^*(\tilde{\alpha}, \gamma^{S^*}(\tilde{\alpha})))} \left( K_{SA} l(Q_d^*(\tilde{\alpha}, \gamma^{S^*}(\tilde{\alpha}))) + K_S \right) = \tilde{\alpha},
\]

since \( \tilde{Q}_d^*(\tilde{\alpha}, \gamma^{S^*}(\tilde{\alpha})) \) satisfies (20) and \( \sum_{i=1}^{n} C^i = C_S + C_{SA} \). We then have that \( \alpha_0 = 1 - \sum_{i=1}^{n} \alpha_i = 1 - \tilde{\alpha} = \tilde{\alpha}_0 \), so that the assembler’s profit margin is the same as in the series system, and \( Q_i^S(\alpha_i) \leq Q_0^I(\tilde{\alpha}_0) = Q_0^I(\alpha_0) \) for \( i = 1, ..., n \). This implies that the vector of prices \( \alpha \) yields, in the general system, an equilibrium capacity of \( Q_d^*(\alpha, \gamma^{S^*}(\alpha)) = Q_d^*(\alpha, \gamma^{S^*}(\tilde{\alpha})) \).

We conclude that the same assembler profit can be achieved in the general system with the vector of prices \( \alpha \). □

**Proof of Proposition 5.** Without loss of generality (by Proposition 4) we restrict attention to a series system with a single supplier and a single subassembler. For any initial cost configuration let \( \alpha_i^* \) be the optimal assembler’s price and let \( Q_d^* \) be the resulting equilibrium system capacity so that \( Q_d^* \) satisfies (12), i.e.,

\[
-(k_{10} + k_1) + (\alpha_1^* - (c_{10} + c_{11} + c_1))\tilde{F}(Q_d^*) - k_{11} l(Q_d^*) = 0. \tag{21}
\]

It is immediately clear from (21) that a shift in unit component production cost between \( c_{10} \) and \( c_{11} \) will have no impact on the equilibrium system capacity, and as a result will have no impact on the assembler’s optimal profit.
Now consider a unit assembly capacity cost shift $\varepsilon > 0$ from supplier to subassembler, i.e., a shift from $k_{11}$ and $k_{10}$ to $k_{11} - \varepsilon$ and $k_{10} + \varepsilon$. Define $\alpha_1 = \alpha_1^* + \frac{\varepsilon}{F(Q_d^*)} (1 - l(Q_d^*)) \leq \alpha_1^*$, since $l(Q_d^*) \geq 1$. If the assembler responds with the new price $\alpha_1$, then $Q_d^*$ is the solution to the first order condition

$$-(k_{10} + k_1 + \varepsilon) + (\alpha_1 - (c_{10} + c_{11} + c_1)) \overline{F}(Q) - (k_{11} - \varepsilon) l(Q) = 0.$$ 

Substituting this same system capacity, along with the new price, into the assembler’s profit function, we have that

$$\pi_0(\alpha_1|\varepsilon) = -k_0 Q_d^* + \left(\alpha_0^* - \frac{\varepsilon}{F(Q_d^*)} (1 - l(Q_d^*)) - c_0\right) \int_0^{Q_d^*} \overline{F}(x) dx$$

$$\geq -k_0 Q_d^* + (\alpha_0^* - c_0) \int_0^{Q_d^*} \overline{F}(x) dx = \pi_0(\alpha_1|0).$$

Since this is just one feasible price response, the assembler’s optimal profit will be at least as high after the cost shift. ■

**Proof of Proposition 6.** Without loss of generality (by Proposition 4) we restrict attention to a series system with a single supplier and a single subassembler. For any initial cost configuration let $\alpha_1^*$ be the optimal assembler’s price and let $Q_d^*$ be the resulting equilibrium system capacity so that $Q_d^*$ satisfies (12), i.e.,

$$-(k_{10} + k_1) + (\alpha_1^* - (c_{10} + c_{11} + c_1)) \overline{F}(Q_d^*) - k_{11} l(Q_d^*) = 0.$$  

First consider a unit assembly cost shift $\varepsilon$ (positive or negative) from subassembler to assembler, i.e., a shift from $c_1$ and $c_0$ to $c_1 - \varepsilon$ and $c_0 + \varepsilon$. If the assembler responds with a new price $\alpha_1 = \alpha_1^* - \varepsilon$, we have $\alpha_1 - (c_{10} + c_{11} + c_1 - \varepsilon) = \alpha_1^* - (c_{10} + c_{11} + c_1)$, so (22) defines the equilibrium system capacity for this modified system as well. Since the system capacity is the same and the assembler faces the same margin $\alpha_0 - (c_0 + \varepsilon) = \alpha_0^* + \varepsilon - (c_0 + \varepsilon) = \alpha_0^* - c_0$, the assembler achieves the same profit as in the original system. Since this is just one feasible price response, the assembler’s optimal profit will be at least as high after the cost shift. A similar argument shows that given the optimal payment from the assembler in the case where a unit assembly cost $\varepsilon$ is shifted, one can construct a corresponding payment in the original system that leads to the same profit for the assembler. This proves the result for $c$.

Now consider a unit assembly capacity cost shift $\varepsilon > 0$ from subassembler to assembler, i.e., a shift from $k_1$ and $k_0$ to $k_1 - \varepsilon$ and $k_0 + \varepsilon$. If the assembler responds with a new price $\alpha_1 = \alpha_1^* - \frac{\varepsilon}{F(Q_d^*)}$, then (22) defines the equilibrium system capacity for this modified system.
as well. Substituting this same system capacity, along with the new cost and price, into the assembler’s profit function yields

\[
\pi_0(\alpha_1|\varepsilon) = -(k_0 + \varepsilon)Q_d^* + \left(\alpha_0^* + \frac{\varepsilon}{F(Q_d^*)} - c_0\right) \int_0^{Q_d^*} F(x) dx \\
\geq -k_0Q_d^* + (\alpha_0^* - c_0) \int_0^{Q_d^*} F(x) dx - \varepsilon Q_d^* + \frac{\varepsilon}{F(Q_d^*)} Q_d^* F(Q_d^*) \\
= -k_0Q_d^* + (\alpha_0^* - c_0) \int_0^{Q_d^*} F(x) dx,
\]

which is the assembler’s optimal profit before the cost and price shift. Since this is just one feasible price response, the assembler’s optimal profit will be at least as high after the cost shift. Interpreting this in the reverse direction establishes the result for \(k\). ■

**Proof of Theorem 2.** Following Proposition 4 and Remark 1, we will consider traditional and modular assembly systems in their pure-series form. We will compare the assembler’s expected profit under a traditional system to his profit under a modular system in which he outsources unit assembly costs \(c_\Delta\) and assembly capacity costs \(k_\Delta\). That is, the traditional system has one supplier with costs \(c_{10} + c_{11}\) and \(k_{10} + k_{11}\) (or, equivalently, two suppliers with costs \(c_{10}\) and \(k_{10}\), and \(c_{11}\) and \(k_{11}\), respectively) and the assembler has costs \(c_0\) and \(k_0\), while the modular system has a single supplier with costs \(c_{11}\) and \(k_{11}\), a subassembler with costs \(c_{10} + c_\Delta\) and \(k_{10} + k_\Delta\), and an assembler with costs \(c_0 - c_\Delta\) and \(k_0 - k_\Delta\).

The superscript \(m\) denotes the modular assembly system, and the superscript \(t\) denotes the traditional system. Also, define \(C = c_0 + c_{10} + c_{11}\) and \(K = k_0 + k_{10} + k_{11}\).

Take \(\alpha_1^{*m}\) to be the optimal assembler payment in the modular assembly system and \(Q^{*m}\) the system’s equilibrium capacity. Observe from (15) that \(\pi_0^m(\alpha_1^{max}) = -(k_0 - k_\Delta)Q_0^m(\alpha_0(\alpha_1^{max})) + \left(\alpha_0(\alpha_1^{max}) - (c_0 - c_\Delta)\right)E[\min\{Q_0^m(\alpha_0(\alpha_1^{max})), D]\} = \max_{Q}\left\{- (k_0 - k_\Delta)Q + \left(\alpha_0(\alpha_1^{max}) - (c_0 - c_\Delta)\right)E[\min\{Q, D]\} \right\} > 0 = \pi_0(\alpha_1^{min})\), since \(\alpha_0(\alpha_1^{max}) > k_0 - k_\Delta + c_0 - c_\Delta\). This implies that \(\alpha_1^{*m} > \alpha_1^{min} = c_{10} + c_{11} + k_{10} + k_{11}\), and the optimal expected profit for the assembler is \(\pi_0^m(\alpha_1^{*m}) = -(k_0 - k_\Delta)Q^{*m} + (1 - \alpha_1^{*m} - (c_0 - c_\Delta))E[\min\{Q^{*m}, D]\}].

In general, for any \(\alpha_1\) chosen by the assembler, we have from (12) that \(\frac{\partial \pi_0^m(Q|\alpha_1)}{\partial Q} = -(k_{10} + k_\Delta) + (\alpha_1 - c_{10} - c_{11} - c_\Delta)F(Q) - k_{11} I(Q) = 0\). (23)

From the concavity of \(\pi_0^m(Q|\alpha_1)\), which follows from (11), we have that (23) has a unique solution for any value of \(\alpha_1\). This allows us to represent

\[
\alpha_1(Q) = c_{10} + c_\Delta + c_{11} + \frac{k_{10} + k_\Delta + k_{11} I(Q)}{F(Q)},
\]

34
so that $Q_m$ maximizes $\pi_0^m(Q) = -(k_0 - k_\Delta)Q + \left(1 - C - \frac{k_{10} + k_{11}l(Q)}{F(Q)}\right)E[\min\{Q, D\}]$. In particular,

$$\frac{\partial \pi_0^m(Q_m)}{\partial Q} = -(k_0 - k_\Delta) + (1 - C)F(Q_m) - (k_{10} + k_\Delta + k_{11})l(Q_m) + k_{11}l(Q_m)(1 - l(Q_m)) - \frac{k_{11}l(Q_m)}{F(Q_m)}\int_0^{Q_m} F(x)dx = 0. \tag{25}$$

For the traditional system, we know that given a payment $\alpha_1$ by the assembler, the supplier will set a capacity level given by

$$Q^f(\alpha_1) = \frac{\overline{Q}}{\alpha_1 - c_{10} - c_{11}}, \tag{26}$$

which leads to assembler’s expected profit given by $\pi^f_0(\alpha_1) = -k_0Q^f(\alpha_1) + (1 - \alpha_1 - c_0)E[\min\{Q^f(\alpha_1), D\}]$.

Replacing $Q^f(\alpha_1)$ in the left-hand side of (23), we get that

$$\frac{\partial \pi^m_1(Q^f(\alpha_1^m)|\alpha_1)}{\partial Q} = k_{11} \left(1 - l(Q^f(\alpha_1^m))\right) - k_\Delta - c_\Delta\overline{F}(Q^f(\alpha_1^m)) < 0,$n

so that $Q^m < Q^f(\alpha_1^m)$. Since $Q^f(\alpha_1)$ is increasing in $\alpha_1$, let $\tilde{\alpha}_1 < \alpha_1^m$ be such that $Q^m = Q^f(\tilde{\alpha}_1)$. More specifically, from (24) and (26) we have that

$$\alpha_1^m = \tilde{\alpha}_1 + c_\Delta + \frac{k_\Delta + k_{11}l(Q^f(\tilde{\alpha}_1) - 1)}{\overline{F}(Q^f(\tilde{\alpha}_1))} > \tilde{\alpha}_1 + c_\Delta + \frac{k_\Delta}{\overline{F}(Q^f(\tilde{\alpha}_1))}. \tag{27}$$

Thus,

$$\pi^m_0(\alpha_1^m) = -(k_0 - k_\Delta)Q^f(\tilde{\alpha}_1) + (1 - \alpha_1^m - (c_0 - c_\Delta))E[\min\{Q^f(\tilde{\alpha}_1), D\}]$$

$$< -(k_0 - k_\Delta)Q^f(\tilde{\alpha}_1) + \left(1 - \tilde{\alpha}_1 - c_0 - \frac{k_\Delta}{\overline{F}(Q^f(\tilde{\alpha}_1))}\right)E[\min\{Q^f(\tilde{\alpha}_1), D\}]$$

$$= -(k_0 - k_\Delta)Q^f(\tilde{\alpha}_1) - k_\Delta\frac{E[\min\{Q^f(\tilde{\alpha}_1), D\}]}{\overline{F}(Q^f(\tilde{\alpha}_1))} + (1 - \tilde{\alpha}_1 - c_0)E[\min\{Q^f(\tilde{\alpha}_1), D\}]$$

$$\leq -(k_0Q^f(\tilde{\alpha}_1) + (1 - \tilde{\alpha}_1 - c_0)E[\min\{Q(\tilde{\alpha}_1), D\}) = \pi^f_0(\tilde{\alpha}_1),$$

where the first inequality follows from (27) and the second one from the fact that $E[\min\{Q, D\}] > Q\overline{F}(Q)$. The amount $\tilde{\alpha}_1$ is one feasible payment for the traditional system, so that the assembler’s expected profit in the traditional system is higher than his profit in the modular system.

From the expression in (26), we can write

$$\pi^f_0(Q) = -k_0Q + \left(1 - C - \frac{k_{10} + k_{11}}{\overline{F}(Q)}\right)E[\min\{Q, D\}],$$

35
which is again concave in $Q$ by (11). Its unique maximum $Q^{t*}$ solves

$$\frac{\partial \pi_0^t}{\partial Q} = -k_0 + (1 - C)F(Q) - (k_{10} + k_{11})l(Q) = 0. \quad (28)$$

Since $l(Q) > 1$ and $l'(Q) > 0$, we have from (25) and the first equality of (28) that

$$\frac{\partial \pi_0^t(Q^{sm})}{\partial Q} = k_\Delta (l(Q^{sm}) - 1) - k_{11}l(Q^{sm}) (1 - l(Q^{sm})) + \frac{k_{11}l'(Q^{sm})}{F(Q^{sm})} \int_0^{Q^{sm}} F(x) dx > 0,$$

implying that $Q^{sm} < Q^{*t}$.

Finally, let $Q^C = \mathcal{F}^{-1}\left(\frac{K}{1-C}\right)$ be the centralized optimal capacity for the system, i.e., the maximum of the concave profit function $\pi^C(Q) = -KQ + (1 - C)E[\min\{Q, D\}]$. Note that $\frac{\partial \pi_0^t(Q^C)}{\partial Q} = (k_{10} + k_{11})(1 - l(Q^C)) < 0$, which implies that $Q^{sm} < Q^{*t} < Q^C$. From the concavity of $\pi^C$, we conclude that $\pi^C(Q^{sm}) < \pi^C(Q^{*t}) < \pi^C(Q^C)$. \hfill \lrcorner

**Proof of Proposition 7.** (a) Since all the cost parameters are fixed, regardless of how much of the assembly work is outsourced under modular assembly, the traditional system has an optimal capacity level $Q^{*t}$ and optimal profit for the assembler $\pi_0^t(Q^{*t})$.

For the modular assembly system, and for any given amount of assembly $k_\Delta$ outsourced to the subassembler, consider the extreme case where the subassembler can process that outsourced assembly work for free. In this case, it follows from (12) that

$$\alpha_1(Q) = C + \frac{k_{10} + k_{11}l(Q)}{F(Q)}, \quad (29)$$

where $C = c_{10} + c_{11} + c_0$. Then, the assembler’s profit, for a given $k_\Delta$, as a function of the capacity level $Q$ is:

$$\pi^m_0(Q|k_\Delta) = -(k_0 - k_\Delta)Q + \left(1 - C - \frac{k_{10} + k_{11}l(Q)}{F(Q)}\right)E[\min\{Q, D\}]. \quad (30)$$

Let $Q^{sm}(k_\Delta)$ denote the capacity level that maximizes (30). Note that if $\pi^m_0(Q^{sm}(k_\Delta)|k_\Delta) < \pi_0^t(Q^{*t})$, then the expected assembler’s profit under modular assembly is lower than his profit under traditional assembly if the assembler outsources an amount $k_\Delta$ of his assembly cost and even if the subassembler could do that work for free. If, on the other hand, $\pi^m_0(Q^{sm}(k_\Delta)|k_\Delta) > \pi_0^t(Q^{*t})$, then there exists a unit capacity cost $0 < k^*(k_\Delta) \leq k_\Delta$ such that break-even is achieved if the subassembler’s unit capacity cost is $k^*(k_\Delta)$. Finally, if $\pi^m_0(Q^{sm}(k_\Delta)|k_\Delta) = \pi_0^t(Q^{*t})$, then break-even is achieved with $k^*(k_\Delta) = 0$.

Since $0 \leq k_\Delta \leq k_0$, let us investigate what happens at both ends of this interval. For $k_\Delta = 0$, the assembler is not outsourcing any work to the subassembler and it follows from (30) and
Theorem 2 that \( \pi_0^m (Q^{*m}(k_\Delta = 0)|k_\Delta = 0) < \pi_0^t (Q^{*t}) \). Clearly, \( \pi_0^m (Q^{*m}(k_\Delta)|k_\Delta) \) is strictly increasing in \( k_\Delta \). Thus, if \( \pi_0^m (Q^{*m}(k_\Delta = k_0)|k_\Delta = k_0) < \pi_0^t (Q^{*t}) \), then we are in the situation described in (a)(ii). If \( \pi_0^m (Q^{*m}(k_\Delta = k_0)|k_\Delta = k_0) \geq \pi_0^t (Q^{*t}) \), then \( \pi_0^m (Q^{*m}(k_\Delta)|k_\Delta) \) crosses \( \pi_0^t (Q^{*t}) \) only once at a point \( \hat{k}_\Delta \). In this case, (a)(i) holds, proving part (a).

(b) Consider now the function

\[
\hat{\pi}_0^m (Q|k_\Delta) = -(k_0 - k_\Delta)Q + \left(1 - \frac{k_{10} + k^*(k_\Delta) + k_{11}l(Q)}{F(Q)}\right) E[\min\{Q, D\}].
\] (31)

for \( k_\Delta \geq \hat{k}_\Delta \). This function represents the assembler profit when he passes costs \( k_\Delta \) to the subassembler, and the subassembler reduces those costs to exactly \( k^*(k_\Delta) \). Let \( \hat{Q}^{*m}(k_\Delta) \) denote the capacity that maximizes (31). By construction, \( \hat{\pi}_0^m (\hat{Q}^{*m}(k_\Delta)|k_\Delta) = \pi_0^t (Q^{*t}) \), a constant with respect to \( k_\Delta \). We then have

\[
0 = \frac{d\hat{\pi}_0^m (\hat{Q}^{*m}(k_\Delta)|k_\Delta)}{dk_\Delta} = \frac{\partial \hat{\pi}_0^m}{\partial k_\Delta} + \frac{\partial \hat{\pi}_0^m}{\partial k^*} \frac{\partial k^*}{\partial k_\Delta} + \frac{\partial \hat{\pi}_0^m}{\partial \hat{Q}^{*m}} \frac{\partial \hat{Q}^{*m}}{\partial k_\Delta} = \hat{Q}^{*m}(k_\Delta) - \frac{E[\min\{\hat{Q}^{*m}(k_\Delta), D\}]}{F(\hat{Q}^{*m}(k_\Delta))} \frac{\partial k^*}{\partial k_\Delta},
\]

since \( \frac{\partial \hat{\pi}_0^m}{\partial Q}(\hat{Q}^{*m}(k_\Delta)|k_\Delta) = 0 \), by the first order conditions. We conclude that

\[
0 < \frac{\partial k^*}{\partial k_\Delta} = \frac{\hat{Q}^{*m}(k_\Delta)F(\hat{Q}^{*m}(k_\Delta))}{E[\min\{\hat{Q}^{*m}(k_\Delta), D\}]} < 1,
\] (32)

where the second inequality holds since \( E[\min\{Q, D\}] > Q^\top (Q) \). This implies that both \( k^*(k_\Delta) \) and \( k_\Delta - k^*(k_\Delta) \) are increasing in \( k_\Delta \). Finally, \( (k_\Delta - k^*(k_\Delta)) / k_\Delta \) is decreasing in \( k_\Delta \) if and only if \( k^*(k_\Delta) / k_\Delta \) increases with \( k_\Delta \) if and only if \( \frac{\partial k^*}{\partial k_\Delta} k_\Delta \geq k^*(k_\Delta) \). By the equality in (32), the latter holds if and only if

\[
y(k_\Delta) \overset{\text{def}}{=} k_\Delta \hat{Q}^{*m}(k_\Delta) - k^*(k_\Delta) \frac{E[\min\{\hat{Q}^{*m}(k_\Delta), D\}]}{F(\hat{Q}^{*m}(k_\Delta))} \geq 0.
\] (33)

Note that \( \pi_0^t (Q^{*t}) = \hat{\pi}_0^m (\hat{Q}^{*m}(k_\Delta)|k_\Delta) \)

\[
\begin{align*}
&= \left\{-k_0 \hat{Q}^{*m}(k_\Delta) + \left(1 - \frac{k_{10} + k_{11}l(\hat{Q}^{*m}(k_\Delta))}{F(\hat{Q}^{*m}(k_\Delta))}\right) E[\min\{\hat{Q}^{*m}(k_\Delta), D\}] \right\} + y(k_\Delta) \\
&\leq \pi_0^m (Q^{*m}(k_\Delta = 0)|k_\Delta = 0) + y(k_\Delta) \leq \pi_0^t (Q^{*t}) + y(k_\Delta),
\end{align*}
\]

where the first inequality follows from the optimality of \( Q^{*m}(k_\Delta = 0) \) when \( k_\Delta = 0 \) in (30) and the second inequality follows from Theorem 2. This shows (33) and completes the proof of part (b).

(c) The fact that \( c_\Delta - c^*(c_\Delta) \) remains constant is a direct consequence of Proposition 6. Suppose \( c^*(c_\Delta) \) achieves break even for a given \( c_\Delta \), and consider an increase in \( c_\Delta \). If
we increase $c_\Delta$ by $\varepsilon$, this corresponds to simply shifting cost (with no additional reduction) from the assembler to the subassembler. Proposition 6 says that such a shift does not change assembler profit. As a result, $c^*(c_\Delta) + \varepsilon$ achieves break even for $c_\Delta + \varepsilon$ — i.e., no additional absolute cost reduction is required as $c_\Delta$ increases. It also follows that $(c_\Delta - c^*(c_\Delta))/c_\Delta$ is decreasing in $c_\Delta$. ■