Time Value of Commercial Product Returns

V. Daniel R. Guide Jr.
Smeal College of Business, The Pennsylvania State University, University Park, Pennsylvania 16802, dguide@psu.edu

Gilvan C. Souza
Robert H. Smith School of Business, University of Maryland, College Park, Maryland 20742,

Luk N. Van Wassenhove
INSEAD, Boulevard de Constance, 77305 Fontainebleau, France,

Joseph D. Blackburn
Owen School of Management, Vanderbilt University, Nashville, Tennessee 37203,

Manufacturers and their distributors must cope with an increased flow of returned products from their customers. The value of commercial product returns, which we define as products returned for any reason within 90 days of sale, now exceeds U.S. $100 billion annually (Stock et al. 2002). Although the reverse supply chain of returned products represents a sizeable flow of potentially recoverable assets, only a small fraction of the value is currently extracted by manufacturers; a large proportion of the product value erodes away because of long processing delays. Thus, there are significant opportunities to build competitive advantage from making the appropriate reverse supply chain design choices. In this paper, we present a network flow with delay models that includes the marginal value of time to identify the drivers of reverse supply chain design. We illustrate our approach with specific examples from two companies in different industries and then examine how industry clockspeed generally affects the choice between an efficient and a responsive returns network.

Keywords:

History: Accepted by William S. Lovejoy, operations and supply chain management; received June 23, 2003.

This paper was with the authors 1 year and 2 months for 3 revisions.

1. Introduction
Manufacturers and their distributors must cope with an increased flow of returned products from their customers. The value of commercial product returns, which we define as products returned for any reason within 90 days of sale, now exceeds U.S. $100 billion annually (Stock et al. 2002). Although the reverse supply chain of returned products represents a sizeable flow of potentially recoverable assets, only a small fraction of the value is currently extracted by manufacturers. A large proportion of the product value erodes away in the returns process. Most returns processes in place today were developed for an earlier environment in which return rates were low and the value of the asset stream was insignificant. Returns processes were typically designed for cost efficiency where collection networks minimized logistics costs and the need for managerial oversight. For example, Stock et al. (2002) describe Sears’ cost-effective transportation network serving three central returns processing centers.

Although cost-efficient logistics processes may be desirable for collection and disposal of products when return rates are low and profit margins are comfortable, this approach can actually limit a firm’s profitability in today’s business environment. The design of processes driven by a narrow operational cost focus can create time delays that limit the options available for reuse. These limited product disposition options can lead to substantial losses in product value recovery. This is typically the case for short-life-cycle, time-sensitive products where these losses can exceed 30% of product value. There is a need for design strategies for product returns that emphasize asset recovery in addition to operating costs, and that need motivates this research.

We consider the problem of how to design and manage the reverse supply chain to maximize net asset value recovered from the flow of returned products. Unlike forward supply chains, no principles of design strategy for returns processing have been established. Blackburn et al. (2004) hypothesize that the marginal value of time can be used to help managers design the right reverse supply chain. Their hypotheses are supported by case studies of several reverse supply chains. We evaluate alternative reverse supply chain designs using network flow models capturing the effects of delays on costs and revenues. Our alternative network designs are derived from two sources: (1) observations of emerging practices
in returns processing and (2) the research on design strategies for forward supply chains.

Our models are built and validated using data collected through in-depth studies of the returns processes at Hewlett-Packard Company (HP) and Robert Bosch Tool Corporation (Bosch). These two firms' product return environments exhibit significant differences in processing and delay costs, and we show that these should lead to alternative network designs, offering useful insights into what drives these decisions. We subsequently use these two cases as a basis for sensitivity analysis and test the generality of our insights.

This paper is organized as follows. In §2, we review the relevant literature. In §3, we present an overview of the product returns system for two manufacturers, HP and Bosch, which serves as a motivation for the model. In §4, we present the model and theoretical results. In §5, we study ways to improve network responsiveness. In §6, we analyze a partially decentralized network for handling product returns. In §7, we apply the results to HP and Bosch, using empirical data from these manufacturers. Finally, we conclude in §8.

2. Literature Review

Although manufacturers have a growing interest in extracting value from commercial product returns, there has been little research on how to design the reverse supply chain for this purpose. However, extensive research has been conducted on managing product return flows for the recovery of products at their end of use (EOU) or end of life (EOL), where products are prevented from entering the waste stream via value and materials recovery systems. Fleischmann (2001), Guide (2000), and Guide and Van Wassenhove (2003) offer comprehensive reviews of the remanufacturing, reverse logistics, and closed-loop supply chain research on EOU/EOL returns processes. Most of these studies focus on cost-efficient recovery and meeting environmental standards, or both. This literature has focused on operating issues (e.g., inventory control, scheduling, materials planning) and the logistics of product recovery. Few papers take a business perspective of how to make product returns operations profitable (see Guide and Van Wassenhove 2001 for a discussion and Guide et al. 2003 for a modeling example).

Much of the previous research on commercial product returns documents the return rates of different product categories and the cost of processing returns. This research finds that return rates vary widely by product category, by season, and across global markets. For example, product return percentages can vary from 5%–9% for hard goods and up to 35% for high fashion apparel. Return percentages are also typically much higher for Internet and catalogue sales. Other research has found that, because of differences in customer attitudes and retailers' return policies, the proportion of returned product tends to be considerably higher in North America. Many retailers in the United States permit returns for any reason within several months of sale. Return policies have been much more restrictive in Europe and, consequently, return rates were markedly lower. However, return rates are rising in Europe rapidly because of new European Union policies governing Internet sales and the entry of powerful U.S.-based resellers. Additionally, companies have seen an increase in commercial returns disguised as defects from large resellers in the United Kingdom (Helbig 2002). Recent studies reported in the trade literature also reveal that returns may cost as much as three to four times the cost of outbound shipments (Andel and Aichlmayr 2002). Although these reports have raised management's awareness of the problem of product returns, the issue of how to extract more value from the returns stream has been largely ignored.

From a marketing perspective, research examines how returns policies affect consumer purchase probability and return rates. Wood (2001) found that more lenient policies tended to increase product returns, but that the increase in sales was sufficient to create a positive net sales effect. Other research has focused on the problem of setting returns policy between a manufacturer and a reseller and the use of incentives to control the returns flow (Padmanabhan and Png 1995, 1997; Pasternack 1985; Davis et al. 1995; Tsay 2001). Choi et al. (2004) study the effect of an e-marketplace on returns policy in which Internet auctions are used to recover value from the stream of product returns.

Supply Chain Design Strategy. A number of researchers have contributed to the development of design strategy for forward supply chains and our models are motivated by this work (Swaminathan and Tayur 2003, Fisher 1997, Lee and Whang 1999, Lee and Tang 1997, Feitzinger and Lee 1997). We are able to confirm a set of design principles for reverse supply chains. Fisher (1997) recommends (cost-)efficient supply chains for functional products (low demand uncertainty), and responsive supply chains for innovative products (high demand uncertainty). We observe that a (cost-)efficient returns network equates to a centralized structure and a responsive network equates to a decentralized one; we relate products with high time-value decay to Fisher's (1997) innovative products. However, we find that in reverse supply chain design, it is early, not delayed, product differentiation that determines profitability.
Valuing Time in Supply Chains. A significant difference between our model and previous research on reverse supply chains is that we explicitly capture the cost of lost product value because of time delays at each stage of the returns process. Studies of time-based competition (Blackburn 1991) have demonstrated that faster response in business processes can be a source of competitive advantage, and other studies have shown how to quantify the effect of time delays in traditional make-to-stock supply chains (Blackburn 2001). In his book *Clockspeed*, Fine (1998) shows that the effects of speed vary across industries and product categories, and he uses these concepts to link supply chain strategies to product architecture. This earlier work provides the motivation for our models that specifically incorporate the cost of time delays and its effect on asset recovery.

3. Commercial Returns at HP and Bosch

Customers may return products for a variety of reasons (see Tables 1 and 2), many of which may be classified as nondefective. Some of these nondefective returns are *new returns*, because they are essentially unused products that may be resold after visual inspection and repackaging. HP estimates the cost of product returns at 2% of total outbound sales for North America alone (Davey 2001). Figure 1 shows the flow for product returns in generic terms.

3.1. Case 1: Hewlett-Packard Inkjet Printers

HP’s product returns strategy is focused on recovering maximum value from the returns and developing capabilities that would put HP in a position of competitive advantage. HP’s inkjet printer division handled more than 50,000 returns per month in North America in 1999 (Davey 2001). The most recent trend estimates show a 20% increase. Inkjet printers have a relatively short life cycle, with a new model being introduced every 18 months on average.

Table 1 Breakdown of Reasons for Commercial Product Returns of HP Printers

<table>
<thead>
<tr>
<th>Reason for return</th>
<th>Description</th>
<th>Percentage of returns (%)</th>
<th>Procedure after return</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product defective</td>
<td>A truly defective product—it simply does not function as intended</td>
<td>20.0</td>
<td>Product is tested, remanufactured (low or high touch), and sold to a secondary market (sell as remanufactured).</td>
</tr>
<tr>
<td>Could not install</td>
<td>The customer could not install the product correctly. Box opened, but product was never used.</td>
<td>27.5</td>
<td>Product is tested for number of pages printed; if this number is zero, then the product is reboxed and shipped back to the forward distribution center to be sold as new. Otherwise, it is shipped to appropriate remanufacturing facility.</td>
</tr>
<tr>
<td>Performance not compatible with user needs</td>
<td>The product did not meet the user’s needs. Print quality was too low, printing speed was too slow, etc.</td>
<td>40.0</td>
<td></td>
</tr>
<tr>
<td>Convenience returns</td>
<td>The product was returned for a host of reasons (remorse, rental, better price, etc.)</td>
<td>12.5</td>
<td></td>
</tr>
</tbody>
</table>

Products returned to the reseller are stored until they are transported to the central HP returns depot outside Nashville, Tennessee, where credit is issued. No hard data is available on how long the returned products spend waiting for transport at the reseller. This can vary drastically from reseller to reseller, but HP managers believe products could spend as long as four weeks stored in areas where they are “out of sight, out of mind” (Davey 2001).

Inkjet printers are delivered via truck and are unloaded and stored in holding areas at the depot to await disposition. The time required for transportation ranges from 6 to 13 days, depending on the distance to be traveled. The receipt and credit issuance take an average of four days. After credit issuance, returns are sorted by product line. Inkjet printers are tested, evaluated, and sent to one of several facilities. All HP printers have an electronic counter that allows a technician to determine how many copies have been printed.

Presently, the average remanufacturing time is 40 days. All remanufactured HP inkjet printers are sold in secondary markets under the direction of a dedicated sales representative.

3.2. Case 2: Robert Bosch Tool Corporation

Bosch’s Skil line is aimed at the consumer market. These tools are reasonably priced and have small profit margins because of the competitive nature of the market. The current product returns process is a
result of the 90-day returns policy, which is intended to attract customers.

Customers return products directly to resellers. The life cycle of power tools currently averages six years. Table 2 shows the primary reasons customers return products (Wolman 2003). The reseller holds the returned tools in an return-to-vendor (RTV) cage. This inventory is held until a Bosch salesperson is available to perform disposition on the product. The period of time between receipt of product and disposition is again highly variable, depending on the work load of the salesperson, with times ranging from one to four weeks (Valenta 2002). The returned products are sent to Walnut Ridge, Arizona, if a product is deemed to be a straightforward remanufacture and to Addison, Illinois, if the problem appears to be more technical in nature. Products are transported in bulk via trucks to the appropriate remanufacturing facility. Products are diagnosed by technicians and remanufactured when possible. Products are discarded if reconditioning is not possible or likely to be very expensive. The reconditioned products are sold mainly to liquidators at an average of 15% below the retail price for the new product.

4. A Simple Analytical Model for the Time Value of Product Returns

We present an analytical model that computes the value of time in a closed-loop supply chain and provides closed-form expressions that allow a manager to quickly compute the value of reducing delays. In §5, we discuss specific actions aimed at reducing delays in the network. We also developed a simulation model in ARENA that allowed us to confirm the model’s robustness under more complex scenarios such as the presence of batching; we comment on this later.

Empirical evidence gathered at HP and Bosch suggests that the rate of commercial returns follows a curve similar to the product life cycle, shifted to the right in the time axis, with a long steady-state period. Figure 2 shows the returns life cycle for an inkjet printer, which has a typical life cycle of 18 months; the steady-state period varies in length from 7 to 13 months. For Bosch power tools, a typical life cycle is six years, with a steady-state period of five years. In the ramp-up period of the life cycle, most returns are used for warranties (i.e., instead of repairing defective products in the field, the firm uses refurbished products originated from convenience returns to replace these defective products), whereas in the ramp-down period, their primary use is for spare parts, after disassembly (Davey 2001).

We develop a profit-maximization model for the steady-state period of the returns life cycle, because of the high volumes involved, the long time frame, and the primary use of returns in the steady-state period for remanufacturing and sales at a secondary market. We model a closed-loop supply chain as a network flow model, shown in Figure 3, where the notation is defined in Table 3. The facilities in the closed-loop supply chain include factory, distribution center, retailer, customer, evaluating facility for returns, remanufacturing, and the secondary market, where remanufactured products are sold. We represent facilities by nodes, and the flow of products through the
nodes is indicated in Figure 3 and described in detail below. To avoid unnecessary confusion, our notation uses parentheses for grouping terms and square brackets for denoting functions, e.g., $r(1-p)$ denotes $r$ times $(1-p)$, and $c[a]$ denotes $c$ as a function of $a$.

Similarly to Toktay et al. (2000) and for ease of exposition, we consider a single retailer. In §7, we show how the model can be easily extended to multiple retailers when we apply it to HP. Each node $i$ experiences a fixed delay $W_{ui}$, there are also transportation delays $\tau_{ij}$ between each pair of nodes $i$ and $j$ in Figure 3, except to and from the customer.

Time $t = 0$ is defined as the beginning of the steady-state period for returns (sales are already in steady state at that time). Time $t = T$ is the end of steady state for sales and returns (whichever is earlier). Thus all nodes are in steady state for the period of analysis. The factory operates in make-to-order mode; $\lambda + (1-p)\lambda_r$ represents the rate of orders to the factory. Products then flow from node to node as they are processed; the flow rates between each pair of nodes $\lambda_{ij}$ are defined in Figure 3, i.e., $\lambda_{sd} = \lambda + (1-p)\lambda_r$, $\lambda_{ds} = \lambda_{sd} = \lambda + (1-p)\lambda_r$, $\lambda_{cr} = \lambda_{rc} = \lambda_r$, $\lambda_{cm} = \lambda_{mc} = (1-p)\lambda_r$, and $\lambda_{cj} = p\lambda_r$.

Inventory is stored as finished goods at the retailer (and thus the delay $W_{cf}$ before the new product is sold) and at the secondary market node (thus the delay $W_{cm}$ before the remanufactured product is sold).

Consistent with empirical data obtained at HP and Bosch, we assume for both new and remanufactured products exponential price decay functions, i.e.,

$$P[t] = P[0]e^{-\lambda t} \quad \text{and} \quad P_m[t] = P_m[0]e^{-\alpha_m t},$$

and exponential variable cost decay functions, i.e., $v[t] = v[0]e^{-\beta t}$ and $v_m[t] = v_m[0]e^{-\phi_m t}$. The continuous-time decay parameters ($\alpha$ and $\alpha_m$, $\phi$ and $\phi_m$) may or may not be equal. All decay parameters can be viewed as a measure of industry clockspeed (see, e.g., Williams 1992, Mendelson and Pillai 1999).

There are handling costs for processing returns, where $h_i$ is the handling cost per unit at facility $i$ ($i = r$ for retailer and $i = e$ for evaluating facility). Transportation and handling costs are assumed constant over time. This is because the decay in prices and variable costs is primarily related to material and product value erosion, which does not hold for transportation and handling costs. All cash flows are discounted at a continuous discount factor $\beta$, which represents the firm’s opportunity cost of capital (i.e., time value of money).

For tractability, we make one assumption:

**Assumption 4-1. New returns are only returned once.** That is, a new return only goes through the cycle in Figure 3 once.

Assumption 4-1 is a reasonable approximation because the fraction of returns that are returned to the forward supply chain is very small, as we document in the case examples described later.

The sequence of events is as follows (see Figure 3):

- **Time $t$: The factory produces** $\lambda + (1-p)\lambda_r$ units at a per unit cost $v[t]$. These units are shipped to the distributor, where they are joined by $p\lambda_r$, new returns (produced at time $t - W_{loop}$, where $W_{loop}$ is the delay through the loop for the network shown in Figure 3), and then transported to the retailer.
- **Time $t + W_{fs}$:** The retailer sells $\lambda + \lambda_r$ units at a per unit price $P[t + W_{fs}]$. After a sojourn time with the customer, $\lambda$ units are returned to the retailer, where they wait until they are shipped to the evaluating facility for sorting and credit issuance.
- **Time $t + W_{fs} + W_{cm}$:** After sorting, the manufacturer issues a credit of $P[t + W_{fs}]$ (selling price) for each of the $\lambda_r$ returns to the retailer. New returns $p\lambda_r$ are shipped to the forward distribution center; non-new returns $(1-p)\lambda_r$ are shipped to the remanufacturing facility.
- **Time $t + W_{fs} + W_{cm}$**: Non-new returns $(1-p)\lambda_r$ are remanufactured at a per unit cost $v_m[t + W_{fs} + W_{cm}]$, and then shipped to the secondary market.
- **Time $t + W_{fs} + W_{cm}$**: Remanufactured products are sold at the secondary market at a per unit price $P_m[t + W_{fs} + W_{cm}]$.

The profit rate at time $t$ for the existing network is

$$\pi[t] = (\lambda + \lambda_r)P[t + W_{fs}] - (\lambda + (1-p)\lambda_r)v[t] - \lambda_pP[t + W_{fs}]e^{-\beta W_{ce}} - p\lambda_r(v[t - W_{loop}] - v[t])$$
tractability, an approximation:

$$+(1-p)\lambda_i(P_m[t + W_{fs} + W_{cz}] - v_m[t + W_{fs} + W_{cm}]) - \sum_{(i,j) \text{ in net}} \lambda_i \tilde{c}_{ij} - \lambda_i \tilde{h}_i - \lambda_j \tilde{h}_j.$$  

(1)

The terms in (1) represent sales revenue for \( \lambda + \lambda \) products sold at a unit price \( P[t + W_{fs}] \) at the retailer, variable production cost at the factory issued at time \( t \), credit issued for \( \lambda \) returns, \( W_{fs} \) time units after they were sold at time \( t + W_{fs} \), difference in variable costs for new returns (i.e., new returns were produced at \( W_{floop} \) time units before other nonreturned products, and hence at a higher cost), unit margin for remanufactured products (unit price \( P_m[t + W_{fs} + W_{cz}] \) minus unit production cost \( v_m[t + W_{fs} + W_{cm}] \)), sum of transportation costs across all network arcs, handling costs at the evaluating facility and retailer, respectively.

The total discounted profit over the steady-state period is \( \Pi = \int_0^T \pi(t)e^{-\beta t} dt \), resulting in

$$\Pi = \lambda(\tilde{P} - \tilde{v}) + \lambda, \tilde{P}e^{-\alpha W_{fs}}(1 - e^{-\beta W_{fs}}) - p\tilde{v}\lambda, e^{-\alpha W_{fs}}(1 - (1 - p)\lambda_i) \cdot \tilde{P}e^{-\alpha(W_{fs} + W_{cz})} - \tilde{v}_m e^{-\phi_m(W_{fs} + W_{cm})} - \tilde{v} - \sum_{(i,j) \text{ in net}} \lambda_i \tilde{c}_{ij} - \lambda_i \tilde{h}_i - \lambda_j \tilde{h}_j,$$

(2)

where, for notational convenience, we define \( \tilde{P} = P[0](1 - e^{-(\alpha + \beta)T})/(\alpha + \beta), \)

$$\tilde{v} = v[0](1 - e^{-(\delta + \beta)T})/((\delta + \beta), \)

$$\tilde{v}_m = v_m[0](1 - e^{-(\phi_m + \beta)T})/((\phi_m + \beta), \)

$$\tilde{P}_m = P_m[0](1 - e^{-(\phi_m + \beta)T})/((\phi_m + \beta), \)

$$\tilde{c}_{ij} = c_{ij}(1 - e^{-\beta T})/\beta_i \text{ and } \tilde{h}_i = h_i(1 - e^{-\beta T})/\beta_i.$$

Thus, \( \tilde{P} \) is the total discounted revenue (including discounting and time-value decay) for the new product over the life cycle \( T \) at a sales rate of one unit per unit time; the other “tilde” parameters are defined similarly.

The terms in (2) represent, discounted over \( T \), the net margin for (net) new products sales (revenues are “discounted” by the delay between production and sale), the “interest” gained by the manufacturer as a result of returns (credit of returns to retailer is issued later than sale), the difference in variable costs for new returns, the margin for remanufactured products, transportation and handling costs.

For the remainder of the analysis, we introduce, for tractability, an approximation:

**Assumption 4-2.** Approximate \( e^{-\alpha W_{fs}} \approx 1 - \alpha W_{fs} \); similar approximations are made for \( e^{\alpha W_{fs}} \), \( e^{\phi_m W_{fs}} \), \( e^{\phi_m W_{fs}} \), and \( e^{\phi_m W_{fs}} \).

Assumption 4-2 is reasonable because for real-life parameters \( \alpha W_{fs} \ll 1 \) (similarly for \( \alpha_m, \phi, \phi_m, \) and \( \beta \))—this approximation implies a maximum error of 0.5% for the numerical examples of §7. We do not use an approximation for \( \tilde{P}, \tilde{v}, \tilde{P}_m, \tilde{c}_{ij}, \tilde{h}_i, \) and \( \tilde{h}_j \), above because \( T \) is considerably larger than any delay \( W_{fs} \) in the network; thus \( \alpha T \gg \alpha W_{fs} \).

Substituting \( W_{floop} = \alpha_1 + \tau_{rd} + W_{fs} \), \( W_{cm} = W_{cm} + \alpha_{cm} + W_{cz} \), and \( W_{cz} = W_{cm} + \alpha_{cm} + W_{cz} \) into (2), and regrouping the terms

$$\Pi \approx \lambda(\tilde{P} - \tilde{v}) + (1 - p)\lambda, \tilde{P}e^{-\alpha W_{fs}}(1 - e^{-\beta W_{fs}}) - \sum_{(i,j) \text{ in net}} \lambda_i \tilde{c}_{ij} - \lambda_i \tilde{h}_i - \lambda_j \tilde{h}_j - (\tau_{rd} + W_{ds})p\lambda, \tilde{v}\phi$$

$$- W_{fs}\{\lambda \tilde{P}\alpha + (1 - p)\lambda, (\tilde{P}_m \alpha_m - \tilde{v}_m \phi_m)\}$$

$$- W_{cv}, \lambda, \{-(\tilde{P}\beta - p\tilde{v}\phi) + (1 - p)(\tilde{P}_m \alpha_m - \tilde{v}_m \phi_m)\} - (\tau_{cm} + W_{mm})(1 - p)\lambda, (\tilde{P}_m \alpha_m - \tilde{v}_m \phi_m) - (\tau_{m2} + W_{m2})(2)(1 - p)\lambda, \tilde{P}_m \alpha_m.$$

(3)

An analysis of (3) allows for an easy visualization for the sources of revenues and costs in the network, as well as the monetary effects of network delays. The first row indicates the steady-state discounted profit without accounting for delays of new and returned products in the network: total discounted new product margins, remanufactured product margins, transportation and handling costs. Equation (3) reveals that this base profit is decreased by the following delays in the network:

(i) The delay of new returns until sale (they are delayed by the loop shown in Figure 3). Thus, a one-day increase in \( \tau_{rd} + W_{ds} \) decreases expected profit by \( p\lambda, \tilde{v}\phi \), corresponding to the daily decrease in total discounted variable production costs. Delays in other components of the loop also affect new products, as explained in (ii).

(ii) The delay of new products to reach the consumer \( W_{fs} \). Thus, a one-day increase in the path between factory and distributor decreases expected profit by \( \lambda \tilde{P}\alpha + (1 - p)\lambda, (\tilde{P}_m \alpha_m - \tilde{v}_m \phi_m) \), corresponding to the daily decrease in total discounted revenues for new and remanufactured products. A one-day increase in the path from distributor to sales decreases expected profit by a higher amount \( \lambda \tilde{P}\alpha + (1 - p) \cdot \lambda, (\tilde{P}_m \alpha_m - \tilde{v}_m \phi_m) + p\lambda, \tilde{v}\phi \) because of its effect on new returns.

(iii) The delay of returned products to reach the evaluating facility \( W_{cz} \). Thus, a one-day increase in the path from consumer to evaluating facility decreases expected profit by \( \lambda, \{-(\tilde{P}\beta - p\tilde{v}\phi) + (1 - p)(\tilde{P}_m \alpha_m - \tilde{v}_m \phi_m)\} \). The time lag for credit issuance to retailers has a positive effect on expected profit. The difference in variable cost for new returns and the daily decrease
in the remanufactured product value have negative effects on expected profit.

(iv) The transportation between the evaluating facility and remanufacturing and remanufacturing delay \( \tau_{em} + W_{mm} \). Thus, a one-day increase in the path from the evaluating facility to remanufacturing decreases expected profit by \( (1 - p)\lambda_r(\tilde{P}_m - \tilde{\nu}_m \phi_m) \), corresponding to the daily decrease in total discounted net revenues for remanufactured products sold in the secondary market.

(v) The delay incurred for transportation and sales in the secondary market \( \tau_{m2} + W_{22} \). Thus, a one-day increase in the path from the remanufacturing facility to the secondary market decreases expected profit by \( (1 - p)\lambda_r\tilde{P}_m\alpha_m \), corresponding to the daily decrease in total discounted sales revenues for remanufactured products sold in the secondary market.

We note that the value of one-day reduction in delays for the reverse network (iii)–(v) depends on the following parameters: return rate \( \lambda_r \), decay parameters for the remanufactured product price \( \alpha_m \) and variable cost \( \phi_m \), proportion of new returns \( p \), remanufactured product revenue \( P_m[0] \) and variable cost \( \nu_m[0] \), variable production cost \( \nu[0] \), and decay parameter for variable production cost \( \phi \) (the term \( \tilde{P} \beta \) is numerically small in our experience). These parameters are all drivers of responsiveness in the reverse network. To gain a better intuition, consider the special case where all value decay parameters are equal (this is the case of HP and Bosch, studied in §7), which we denote by \( \varphi \). Then, the values of one day in the different links of the reverse network (iii)–(v) become \( \lambda_r[(-\tilde{P}\beta - p\tilde{\nu}\varphi) + (1 - p)\varphi(\tilde{P}_m - \tilde{\nu}_m)] \), \( (1 - p)\lambda_r\varphi(\tilde{P}_m - \tilde{\nu}_m) \), and \( (1 - p)\lambda_r\tilde{P}_m\varphi \). In short, ignoring the (numerically small) term \( \tilde{P}\beta - p\tilde{\nu}\varphi \), a day in the reverse network is more valuable if the return rate is higher, fewer new returns are diverted directly into the forward chain, the value decay parameter is higher, the remanufactured product profit margin is higher, and the remanufactured product value is higher. To put it differently, time compression is important in the reverse network for product returns with high recoverable value, high-value decay parameter, and high volume of remanufacturing.

In our simulation model, we examined the impact of batching at the retailer, evaluation, and remanufacturing facilities and observed longer delays and, as a result, greater value decay in products. We also examined the impact of capacity constrained facilities and the results again showed significantly longer delays. These results support the insights gained from the analytical model, and we therefore restrict our attention to the analytical model in the remainder of the paper.

5. Improving Network Responsiveness

The preceding analysis demonstrates the monetary benefits of decreasing delays in different parts of the network. It allows for a time-cost analysis of responsive network designs. In this section, we provide a simple analysis of the optimal level of responsiveness in the network. To provide closed-form expressions, we model the delay at each node by the expected flow time through an \( M/M/1 \) queue, except for the delay at the customer, sales at the retailer, and in the secondary market, where the delay is a constant value. Our choice of \( M/M/1 \) queues for the nodes captures the significant congestion effects observed in practice for the relevant processing facilities and it has been used before in supply chain modeling (e.g., Toktay et al. 2000, Iyer and Jain 2003). It also means that there is no overtaking; that is, all products go through the supply chain on a first-in–first-out mode. We note, however, that other delay expressions are possible (e.g., \( M/M/S \) queue), although they prevent closed-form expressions. Our deterministic flow model with delays now becomes equivalent to an \( M/M/1 \) queuing network model with the expected value substituted for the random flow time in each node to compute the total expected profit over \( T \).

Denoting by \( \mu_i \) the mean processing rate at node \( i \), and using the expressions for expected flow time for an \( M/M/1 \) queue, the expected delays \( W_{ij} \) are computed as follows:

\[
W_{fi} = \frac{1}{\mu_j - (\lambda + (1 - p)\lambda_r)} + \frac{1}{\mu_d - (\lambda + \lambda_r)}
+ \tau_{fd} + \frac{1}{\mu_s}, \quad (4)
\]

\[
W_{ce} = \frac{1}{\mu_c} + \frac{1}{\mu_r - \lambda_r} + \tau_{ce} + \frac{1}{\mu_s - \lambda_r}, \quad (5)
\]

\[
W_{cm} = W_{ce} + \tau_{em} + \frac{1}{\mu_m - (1 - p)\lambda_r}, \quad (6)
\]

\[
W_{c2} = W_{cm} + \tau_{m2} + \frac{1}{\mu_2}, \quad \text{and} \quad (7)
\]

\[
W_{ds} = \frac{1}{\mu_d - (\lambda + \lambda_r)} + \tau_{ds} + \frac{1}{\mu_s}. \quad (8)
\]

After substituting (6)–(8) into (3), we obtain

\[
\Pi \approx \lambda(\tilde{P} - \tilde{\nu}) + (1 - p)\lambda_r(\tilde{P}_m - \tilde{\nu}_m - \tilde{\nu}) - \sum_{(i,j)} \lambda_{ij} \tilde{e}_{ij} - \lambda \tilde{h}_r
- \lambda \tilde{h}_r = \left( \tau_{cd} + \frac{1}{\mu_d - (\lambda + \lambda_r)} + \tau_{ds} + \frac{1}{\mu_s} \right)
\cdot p\lambda_r \tilde{\nu}\varphi - W_{fs}[\lambda \tilde{\nu} + (1 - p)\lambda_r(\tilde{P}_m \alpha_m - \tilde{\nu}_m \Phi_m)]
- W_{ce} \lambda_r[-(\tilde{P} \beta - p \tilde{\nu} \varphi) + (1 - p)(\tilde{P}_m \alpha_m - \tilde{\nu}_m \Phi_m)]
\]
To improve network responsiveness, we can increase \( \mu_i \) at each node (retailer, evaluating, and remanufacturing facilities) and decrease the average transportation times \( \tau_{ij} \) (by colocation of facilities or faster transportation modes). Before analyzing these alternatives, we note that \( \Pi \) is a separable function in each delay variable \( \mu_i \) (that is, \( \partial^2 \Pi / \partial \mu_i \partial \mu_j = 0 \) for \( i \neq j \)) and thus a sufficient condition for (9) to be jointly concave in \( \mu_i \), for all \( i \), is that \( \partial^2 \Pi / \partial \mu_i^2 < 0 \) for all \( i \).

5.1. Increasing Processing Rate of Returns at the Retailers or Evaluating Facilities

Improving responsiveness \( \mu_i \), at the retailer requires investments by the manufacturer according to the unit handling cost \( h_i[\mu_i] \), where we make explicit the dependence of the handling cost with the processing rate. At Bosch, the returns are held at the retailer until a Bosch representative makes a disposition and shipment decision. Bosch can increase the processing rate at each retailer by increasing the number of visits, which may require more service personnel. Similarly, the manufacturer can also improve the processing rate of returns at the central evaluating facility \( \mu_e \). This would again involve investments in workforce for parallel processing or investments in sorting, picking, and routing technology.

To find the optimal level of responsiveness \( \mu^*_i \), we apply the first-order condition to (9), recalling that \( \mu_i^* \), \( i \in \{r, e\} \) impacts \( W_c \), according to (5).

\[
\frac{\partial \Pi}{\partial \mu_i} = 0 \Rightarrow \frac{-(\tilde{P}_B - p\tilde{v}\phi) + (1 - p)(\tilde{P}_m\alpha_m - \tilde{v}_m\phi_m)}{(\mu_i - \lambda_i)^2} = \tilde{h}_i[\mu_i]^*, \quad i \in \{r, e\}.
\]

Sufficient conditions for (3) to be jointly concave (such that the solution to (10) is sufficient for optimality) are that (i) \( \tilde{h}_i[\mu_i] \) be a convex function (including a linear function, which is a reasonable assumption as stated below) and (ii) that \( (1 - p)(\tilde{P}_m\alpha_m - \tilde{v}_m\phi_m) > \tilde{P}_B - p\tilde{v}\phi \); that is, remanufacturing margins are higher than the net (negative) impact of the time lag for returns (i.e., difference between time value of money for credit issuance and production cost lag for new returns) since

\[
\frac{\partial^2 \Pi}{\partial \mu_i^2} = -2\lambda_i \frac{-(\tilde{P}_B - p\tilde{v}\phi) + (1 - p)(\tilde{P}_m\alpha_m - \tilde{v}_m\phi_m)}{(\mu_i - \lambda_i)^3} - \lambda_i \tilde{h}_i'[\mu_i],
\]

which is strictly negative if these two conditions are satisfied.

Now, assume a linear function for the unit handling cost as a function of the processing rate for returns, i.e., \( h_i[\mu_i] = a_i \mu_i + b_i \). This linear function can be justified because return handling operations are labor intensive (Davey 2001). Then, \( \tilde{h}_i[\mu_i] = \tilde{a}_i \mu_i + \tilde{b}_i \), where \( \tilde{a}_i = a_i(1 - e^{-\beta \tau})/\beta \) and a similar expression holds for \( \tilde{b}_i \). For this linear cost case, (10) yields

\[
\mu_i^* = \lambda_i + \sqrt{\frac{(1 - p)(\tilde{P}_m\alpha_m - \tilde{v}_m\phi_m) - (\tilde{P}_B - p\tilde{v}\phi)}{\tilde{a}_i}}, \quad i \in \{r, e\}.
\]

We note that (11) has the solution form of a classic queuing design problem: find the optimal processing rate at an M/M/1 queue that minimizes the expected cost rate (see, e.g., Gross and Harris 1998, p. 304), with waiting cost rate \( (1 - p)(\tilde{P}_m\alpha_m - \tilde{v}_m\phi_m) - (\tilde{P}_B - p\tilde{v}\phi) \) and service cost rate \( \lambda_i \tilde{a}_i \). Only a fraction \( 1 - p \) of all returns \( \lambda_i \) are remanufactured and sold at a revenue of \( \tilde{P}_m \) with an “interest rate” \( \alpha_m \). This revenue is decreased by the total discounted variable remanufacturing costs \( \tilde{v}_m \), which decrease at a rate \( \phi_m \). In addition, the waiting cost rate should be decreased by the time value of money amount corresponding to the daily profit increase of a delayed credit issuance to retailers \( \tilde{P}_B \), but increased by the daily decrease in total variable cost of production for new returns. The optimal return processing rate at either retailer or evaluating facility is not influenced by transportation costs, but it is directly influenced by the remanufactured product margin. Low margins result in designs with a low level of responsiveness. A higher remanufacturing price decay parameter \( \alpha_m \) and a higher variable cost decay parameter \( \phi_m \) (higher clockspeed) increase the waiting cost rate (numerator in the square root of (11)). This increases processing capacity (lowers the waiting time) leading to a more responsive returns network design.

A similar analysis can be conducted for the optimal level of responsiveness in the forward distribution network, i.e., \( \mu_i \), \( i \in \{f, s, d\} \). However, this requires modeling specific costs associated with a level of responsiveness at the factory (increased transportation frequency to the distributor), distributor (more frequent deliveries to retailers) and retailer (advertising, promotion, and pricing), and the focus of this paper is not on forward supply chains.

5.2. Increasing Transportation Responsiveness

Transportation responsiveness in the network can be influenced by design choices such as colocation of facilities or selecting faster transportation. For example, if the firm colocates the remanufacturing and the evaluating facilities, then \( \tau_{cm} = 0 \), and profits increase by \( \tau_{cm}(1 - p)\lambda_i(\tilde{P}_m\alpha_m - \tilde{v}_m\phi_m) \), according to (3).
6. Preponement: Decentralized Returns Network

In this section, we analyze the drivers of alternative structural designs. Figure 3 represents the typical centralized industrial returns evaluation and credit issuance network design where all commercial returns are shipped to a central facility for economies of scale. The benefits in economies of scale for evaluation and credit issuance are clear. Alternatively, consider an innovative design where new returns are sorted and immediately restocked at the retailer. This decentralized design reduces transportation costs, utilization at the central evaluation facility, and consequently, the delay of other returned products. This, in turn, increases their value in the secondary market. We call this decentralized design concept preponement (or early product differentiation) to distinguish it from postponement (or late product differentiation) to distinguish it from preponement.

With preponement, additional work is required at the retailer to handle and repack the returns. Without any capacity adjustment from the existing configuration, the processing rate at the retailer with preponement is evidently lower than in the existing configuration; thus a capacity increase may be warranted. The retailer may need to hire and train workers to perform this task and maintain extra packaging material at the stores. To gain retailer cooperation, the manufacturer may need to offer incentives. Alternatively, the manufacturer could periodically send workers to the retailer’s site to handle the returns, similar to vendor-managed inventory (VMI).

With preponement, there is no need to separate new returns from other returns at the evaluating facility, although the facility still has to issue credit to returns and route them to the appropriate remanufacturing facility. Further, that node experiences a lower flow of products (\(\lambda_c = (1 - p)\lambda_c\), as opposed to \(\lambda_c = \lambda_c\)). Without any capacity adjustment from the existing configuration, the processing rate at the evaluating facility with preponement is evidently higher; thus, a capacity decrease may be attractive.

The decentralized design network is shown in Figure 4. We use a superscript \(p\) to denote, when different, parameters for this proposed preponement network. The flow rates between each pair of nodes are \(\lambda^p_c = p\lambda_c\), \(\lambda^p_e = (1 - p)\lambda_e\), \(\lambda^p_d = \lambda + (1 - p)\lambda_c\), and \(\lambda^p_s = 0\); other flows are as before. As in §4, we do not assume any functional form for the delays at the nodes to keep our results general.

An analysis similar to that performed in §4 provides the total discounted profit over the steady-state period of the life cycle as follows:

\[
\Pi^p \approx \lambda (\tilde{P} - \tilde{v}) + (1 - p)\lambda_c (\tilde{P}_m - \tilde{v}_m - \tilde{v}) - \sum_{(i,j)} \lambda^p_{ij} \tilde{c}_{ij} \\
- \lambda_e \tilde{h}_e - (1 - p)\lambda_c \tilde{h}_c - (W_{cc} + W_{mm} + W_{rr})p\lambda_c \tilde{v} \phi \\
- W_{mm}^p (\tilde{P} \alpha + (1 - p)\lambda_c (\tilde{P}_m \alpha_m - \tilde{v}_m \phi_m)) \\
- W_{cc}^p (1 - p)\lambda_c (\tilde{P}_m \alpha_m - \tilde{v}_m \phi_m) + (W_{cc} + W_{rr})\lambda_e \tilde{P} \beta \\
- (\tau_{em} + W_{mm})(1 - p)\lambda_e (\tilde{P}_m \alpha_m - \tilde{v}_m \phi_m) \\
- (\tau_{m2} + W_{mm})(1 - p)\lambda_e \tilde{P}_m \alpha_m. \\
(12)
\]

We do not include in (12) the incentive, if any, paid by the manufacturer to the retailer, or the extra VMI cost. Our analysis focuses on the total benefits of the proposed network. This benefit can be weighed against these extra monetary incentives or costs. Relative to the centralized network of Figure 3, there are three delays that are different in the preponement network of Figure 4: (1) the delay for the returned product between the consumer and the evaluating facility \(W^p_{cc}\); (2) the delay for the new product between factory and sales \(W^p_{mm}\); and (3) the delay of returns at the retailer \(W^p_{rr}\).

Taking the difference (12)–(3) and defining \(\Delta\) as the difference in delay at node \(i\) between the existing and preponement networks (e.g., \(\Delta_e = W^p_{rr} - W^p_{rr}\)), we state, after some algebra, the following monetary benefits of the proposed decentralized network:

\[
\Pi^p - \Pi = \lambda_e \left\{ (1 - p)(\tilde{P}_m \alpha_m - \tilde{v}_m \phi_m)(\Delta_e + \Delta_s + \Delta) \\
- \tilde{P} \beta (\tau_{ec} + W^p_{ec} + \Delta_e + \Delta_s) \\
+ p \tilde{v} \phi (\tau_{ec} + W^p_{ec} + \Delta_e + \Delta_s + \tau_{ed} + W_{ed} + \tau_{de}) \\
+ p (\tilde{c}_c + \tilde{c}_e + \tilde{c}_d) + (\tilde{h}_e - (1 - p)\tilde{h}^p_{ec}) + (\tilde{h}_r - \tilde{h}^p_{rr}) \right\} \\
+ \lambda \tilde{P} \alpha \Delta_e. \\
(13)
\]
The terms in (13) indicate, respectively:
(i) the increased value obtained from remanufactured products if they reach the secondary market earlier,
(ii) the decrease in profit because there is no time lag for credit issuance for new returns in the preponement network,
(iii) the savings in variable production cost for new returns because they are resold faster,
(iv) the decrease in transportation cost for new returns in the preponement network,
(v) the difference in handling cost at the retailer and evaluating facility, and
(vi) the increased value of new product sales because of reduced delay at the distributor, as a consequence of new returns no longer being routed there.

With the exception of the last term \(\lambda \tilde{P} \alpha \Delta \tau\), which is likely to be small in practice since new returns constitute a small percentage of the flow of products through the distributor in the existing network (\(\Delta \tau\) is a small number), the return rate \(\lambda\), multiplies the entire right-hand side of (13); that is, \(\lambda\), is a scaling parameter for the benefits of preponement. Drivers of the attractiveness of preponement design include, as before, decay parameters for the remanufactured product price \(\alpha_m\) and variable cost \(\phi_m\), the decay rate for variable production cost \(\phi\), proportion of new returns \(p\), the revenue and costs factors of the preponement network, \(P_m[0], v_m[0], v[0]\), transportation, and handling costs (again, the term \(\tilde{P} \beta\) is numerically small in our experience). We develop two general propositions providing insights into three major drivers of attractiveness of the preponement design, i.e., the continuous-time variable production cost decay parameter \(\phi\), the variable cost \(v[0]\), and the proportion of new returns, \(p\).

**Proposition 1.** The benefits of preponement \(\Pi^p - \Pi\) are increasing in \(\phi\) and \(v[0]\) if the time difference for restocking a new return between the existing network (via the evaluating facility) and the preponement network (at the retailer only) is positive; that is,

\[
K = \tau_c + W_{ce} + \tau_{ed} + W_{dd} + \tau_{ds} + \Delta \tau > 0. \tag{14}
\]

**Proof.** The third term of (13) can be written as \(\lambda_p\tilde{v}\phi K\) or

\[
\lambda_p v[0] \frac{1 - e^{-(\phi + \beta)T}}{\phi + \beta} \phi K,
\]

where \(K\) is the term in parenthesis that multiplies \(\tilde{v}\phi K\) in (13). This term, the only in \(\Pi^p - \Pi\) that includes \(\phi\) and \(v[0]\), is increasing in \(\phi\) and \(v[0]\) if \(K > 0\), which results in (14), after we write \(\Delta \tau = W_{ce} - W_{ce}^p\).

It is possible for the benefits of preponement \(\Pi^p - \Pi\) to be positive (or negative) for all meaningful values of \(\phi\); otherwise Proposition 1 implies that there is a \(\phi^*\) such that a decentralized (preponement) network design is preferred if \(\phi \geq \phi^*\); else a centralized network is appropriate. Condition (14) holds for most returns networks because it only requires that all the delays for new returns in the original network exceed the delay at \(\text{only}\) the return node in the preponement network. A similar result can be derived for the other design driver \(p\).

**Proposition 2.** If (14) holds and the total benefits from preponement (time value, transportation, and handling savings) outweigh the potential value that could be gained if new returns were sent to the secondary market (faster with preponement) rather than the primary market, then the benefits of preponement \(\Pi^p - \Pi\) are increasing in \(p\).

**Proof.** Simple algebra shows that the term that multiplies \(p\) in the right-hand side of (13) is

\[
\tilde{v} \phi K \left(\tilde{c}_{ds} + \tilde{c}_{re} + \tilde{c}_{ed} + \tilde{h}_{p}^\prime\right) - (\Delta \tau + \Delta \tau_d + \Delta \tau_e) (P_m \alpha_m - \tilde{v} \phi_m).
\]

The first term is positive because of Proposition 1. The second term represents the transportation and handling savings from preponement. The last term represents the time value gained by selling new returns on the secondary market faster with preponement; that value, however, is foregone because new returns go to the primary market. The last term is the only (potentially) negative term that multiples \(p\) in \(\Pi^p - \Pi\).

Again, it is possible that the benefits of preponement \(\Pi^p - \Pi\) are positive (or negative) for all \(p \in [0, 1]\); otherwise, Proposition 2 implies that there exists a \(p^*\) such that a decentralized network is preferred if \(p \geq p^*\).

Assuming \(M/M/1\) delay expressions at the nodes, and a linear unit handling cost function as before \(h_{p,i}^m[\mu_i] = a_i^m \mu_i^p + b_i^p\), we perform a similar analysis to §5.1 to find the optimal processing capacities at the retailer and evaluating facility. Then,

\[
\mu_r^p = \lambda_r + \sqrt{(1 - p) (P_m \alpha_m - \tilde{v} \phi_m) - (\tilde{P} \beta - p \tilde{v} \phi)/(\tilde{a}^p K)}, \tag{15}
\]

and

\[
\mu_{e}^p = (1 - p) \lambda_e + \sqrt{(P_m \alpha_m - \tilde{v} \phi_m - \tilde{P} \beta)/(\tilde{a}^p e K)}. \tag{16}
\]

It is reasonable to expect that the preponement design option will have higher variable handling costs at the retailer (because of extra tasks) and lower variable handling costs at the evaluation facility (because of less tasks), i.e., \(\tilde{a}^r \geq \tilde{a}\), and \(\tilde{a}^e \leq \tilde{a}^e\). Thus, \(\mu_r^p \leq \mu_r^p\); because (15) only differs from (11) in the denominator inside the square root. Because the handling cost increases linearly with the processing rate at the
retailer, and this rate of increase is higher in the preponement scenario, the optimal processing capacity is smaller in the preponement scenario. Comparing $\mu_r^p$ and $\mu_r^*,$ it is not as straightforward because the lower value of $d_r^p$ tends to increase $\mu_r^p$ relative to $\mu_r^*.$ However, the lower flow of returns $(1 - p)\lambda,$ through the evaluation facility tends to decrease $\mu_r^p$ relative to $\mu_r^*.$ For larger values of $p,$ it is clear that the lower flow effect will tend to dominate (16). In the limit, when $p = 1,$ $\mu_r^p = 0,$ and $\mu_r^p \leq \mu_r^*,$ clearly holds.

In the next section, we apply our theoretical results to HP and Bosch and perform a sensitivity analysis on the key drivers of responsiveness and preponement design alternatives.

7. Application of Model Results
In this section, we apply the theoretical results to actual data from HP and Bosch. The main differences in parameter values for the two firms are product value, life-cycle length, value decay parameters, demand, and return rates. Many of the parameter values are approximately equal for both firms, and for reasons of confidentiality, we use common representative numbers assumed fixed throughout the numerical analysis: a 25% gross margin for new products $(v[0]/P[0] = 0.75),$ a 15% price discount for the remanufactured product relative to the new product $(P_m[0]/P[0] = 0.85),$ and a 5% yearly discount rate $(\beta = 1.4 \times 10^{-4}).$

The price decay parameters for remanufactured and new products are approximately the same $(\alpha = \alpha_m)$ within each company, albeit different between companies. Although different components decay at different rates, we estimate that the overall manufacturing cost of a product decays at a rate roughly equal to the final product’s price decay; that is, $\alpha = \alpha_m = \phi = \phi_m.$ For this reason, we use a single-value decay parameter $\varphi$ for each company. This assumption brings parsimony to the analysis without compromising insights or the order of magnitude of the results. The units of analysis throughout are a full truckload of returned products and a time of one day.

7.1. Hewlett-Packard Inkjet Printers
A delivery truck contains an average of 250 inkjet printers. The median price of an HP inkjet printer is $200, and thus $P[0] = 250 \cdot 200 = 50,000.$ For inkjets, $T = 395$ days (13 months), returns are 5% of net sales, so $\lambda / \Lambda = 0.05.$ The daily return rate averages $\lambda = 6.67$ trucks, and the common-value decay parameter is $\varphi = 1.43 \times 10^{-3}$ (1% per week). The percentage of new returns is $p = 1/3;$ these correspond, in Table 1, to all returns categorized as “could not install,” and a portion of returns categorized as “convenience returns.” The remanufacturing cost is approximately 7.5% of the retail price of a new product; that is, $v_m[0]/P[0] = 0.075.$

Our analysis shows the values of a one-day reduction between different facilities in the returns network: $35,069$ between the evaluating facility and distributor, $93,797$ between the customer and evaluating facility, $72,475$ between the evaluating facility and remanufacturing, and $79,489$ between remanufacturing and the secondary market, respectively. Managers indicate that lead-time reduction in the forward network is currently being pursued at the level of hours, not days. However, opportunities for significantly reducing lead times abound in HP’s reverse supply chain. The sojourn time at retailers, delay between retailers and process completion at the evaluating facility, and delay between the evaluating facility and remanufacturing completion average 10, 8, and 40 days, respectively. We analyze each opportunity separately below.

First, consider the retailer returns processing capacity. For a more realistic analysis, consider multiple retailers. For example, using 1,000 identical retailers with an average sojourn time of 10 days, and assuming $M/M/1$ delays at the retailer implies $1/(\mu_r - \lambda_r/1,000) = 10,$ or a current return processing capacity of $\mu_r = 0.1067.$ If we decrease the average sojourn time by two days (and save approximately $180,000) with the same rate of returns, this implies $\mu_r = 0.1317,$ or a 23% increase in returns processing capacity. To find the optimal processing capacity (11), we require an accurate estimate of handling costs at the retailers.¹

Second, consider transportation to, and sojourn time at, the evaluating facility. Managers at HP believe that this delay can be cut from its current eight days to two days, resulting in a life-cycle savings of approximately $500,000. Finally, the largest opportunity lies in the long delays for shipment from the evaluating facility until completion of the remanufacturing operation, which is currently 40 days. Management believes that a reasonable goal for this delay is 20 days. Achieving this goal implies a life-cycle savings of $1.45$ million. We note that our estimates are conservative, because we do not explicitly account for savings in working capital and the corresponding reduction in inventory holding costs. Thus it appears worthwhile for HP to consider a responsive network design.

We estimate the current discounted life-cycle value of preponement for HP (13) to be roughly $4.0$ million, using the following assumptions: (i) retailers are

¹ We note that the conditions (i) and (ii) for optimality of (11), which are described in the paragraph after (10), are both satisfied. Condition (i) is naturally satisfied because (11) assumes linear handling costs. Condition (ii) is satisfied because $(1 - p)(\bar{P} \alpha_n - \bar{v} \phi_n) = 10,866 > \bar{P} \beta - p \bar{v} \phi = -3,189.$
situated at an average of 1,000 miles from the evaluating facility; (ii) the truckload transportation rate is $1.3$/mile;\(^2\) (iii) the likely increase in handling cost at the retailer is offset by the likely decrease in handling cost at the evaluating facility, and consequently, the difference in total handling costs (across retailer and evaluating facility) between the current and preponement scenarios is negligible; and (iv) the difference in delays between the current and preponement scenarios is negligible (i.e., \(\Delta_d = 0\); \(\Delta_r = 0\)).

Of these $40.0$ million preponement benefits, roughly 20\% are related to the time-value savings in variable costs for new returns (third term in (13)), 82.7\% are related to savings in transportation costs (fourth term in (13)); the second negative term in (13) is small at \(-2.7\%\); the first and last two terms in (13) are zero by our assumptions. It should be clear from these rough-cut calculations that HP has a keen interest in a more detailed analysis of the practical implications of the preponement option. For example, in a more detailed analysis, HP would need to analyze whether to provide financial incentives to the retailer for her to perform preponement, or to do it in “VMI” mode; likely cost increases (to HP) in either case should be investigated and compared against the benefits computed above. Regarding assumption (iii) above, HP may significantly reduce the potential increase in retailer handling costs with preponement by using Galileo, a small device that can be plugged into a printer to immediately reveal the number of pages printed (if the number is zero, the return is considered new and can be reshelved). In general, preponement becomes more attractive if one substitutes technology for labor. Additional investments in technology (e.g., Galileo), if any, should be analyzed also.

7.2. Bosch Power Tools
A delivery truck contains an average of 500 power tools. The average price of a Bosch power tool is $50, and thus \(P[0] = 25,000\). For power tools, \(T = 1,675\) days (55 months). Return rate is 2.6\% of net sales (\(\lambda_r/\lambda = 0.026\), \(\lambda_r = 1.5\), and the common-value decay parameter is \(\phi = 3.5 \times 10^{-4}\) (1\% per month). The percentage of new returns for Bosch is \(p = 0\); Table 2 shows that only 10\% of returns can potentially be unused (category “convenience returns”), however, Bosch indicated that all of its returns have been used to some degree and cannot be considered new. The remanufacturing cost is approximately 7.5\% of the retail price of a new product; that is, \(v_m[0]/P[0] = 0.075\).

The value of reducing one day between the customer and evaluating facility (which is colocated at the new products factory) \(W_{yz}\) is $5,624. The value of one-day reduction between the evaluating facility and remanufacturing and between remanufacturing and the secondary market are $11,623 and $12,748, respectively. Given these results, it appears that Bosch should consider an efficient reverse supply chain network to handle returns.

At Bosch, preponement is a much less viable option than at HP. This is easily explained by the major drivers: a much smaller return rate containing very few new returns and therefore smaller potential transportation cost savings, and a considerably smaller value decay over time, yielding even smaller savings in variable production costs for new returns. Setting up decentralized low-touch remanufacturing facilities (thereby approximating the idea of preponement) would be relatively costly as well, even if all 40\% of nondefective returns (Table 2) could be handled decentrally (and thereby avoid larger transportation costs).

7.3. Sensitivity Analysis
To gain general insights in the drivers of reverse supply chain design, we performed a sensitivity analysis. Using the base numbers for HP’s product value, life-cycle length, and demand volume, we vary the values for the key drivers of reverse supply chain design: the return rate \(\lambda_r\), the common-value decay parameter \(\phi\), the proportion of new returns \(p\), and the remanufactured product profit margin \(P_m[0] - v_m[0]\) (\(P_m[0] = 0.85P[0]\) is fixed, so we vary \(v_m[0]/P[0]\)). We also examine the effect of changes in the life-cycle length \(T\) (demand volume does not impact the reverse network design). We selected the range for these parameters based on representative values for products in a wide range of industries. That is, \(\lambda_r \in [0, 15]\), corresponding to a return rate between 0\% and 12\% of net sales; \(\phi \in [0.0001, 0.004]\), corresponding to monthly value decay rates between 0.3\% and 12\%; \(p \in [0, 0.75]\); \(v_m[0]/P[0] \in [0.025, 0.50]\), corresponding to a remanufacturing cost between 2.5\% and 50\% of the new product price; and \(T \in [180, 1,675]\), corresponding to a life-cycle length between 6 and 55 months. The analyses assume that every unit decrease in returns results in one more unit of net sales; that is, \(\lambda + \lambda_r\) is kept constant at 140 truck-loads per day. We focus on the value of a one-day reduction between the evaluating facility and remanufacturing completion, because this segment has the largest delays. Additional sensitivity analyses were performed for the other segments of the reverse supply chain (e.g., customer and evaluating facility) and the results were similar to the ones discussed here, and therefore omitted.

Figure 5 shows the value of one-day time reduction between the evaluating facility and remanufacturing completion as a function of the return rate \(\lambda_r\) and the

\(^2\)This estimate of transportation rate is based on a U.S. Department of Transportation report http://ops.fhwa.dot.gov/freight/documents/bts.pdf.
time-value decay parameter $\phi$. The marginal value of time becomes important for higher values of the return rate. In those cases, e.g., companies like HP, there are substantial benefits to be gained from considering a responsive reverse supply chain design. Conversely, when return rates are low, a cost-efficient reverse supply chain is favored, even when the marginal value of time is high. Because both returns and the marginal value of time are increasing at a rapid pace globally and across industries, managers need to be aware of the growing potential benefits of adopting a responsive reverse supply chain design.

Figure 6 shows that for companies where both return rates and time-value decay are considerable (using HP’s 5% return rate), the proportion of new returns has a negative linear impact on the value of a one-day time reduction between remanufacturing and sales at the secondary market; this is a result of a lower flow of products that are remanufactured as $p$ increases. However, as mentioned previously, the proportion $p$ is a driver of a decentralized preponement returns network. Figure 7 shows that the value of a one-day reduction increases at a decreasing rate with the life-cycle length $T$; this is, because of discounting and decreasing product values with time. Figure 8 shows the value of one day between evaluating facility and remanufacturing completion as a function of the normalized remanufacturing variable cost; the impact is more significant at high-value decay rates.

Finally, Figure 9 shows the value of preponement as a function of the proportion of new returns. The preponement value is more sensitive to $p$ because approximately 80% of the value of preponement for HP is derived from savings in transportation costs for new returns as compared to 20% derived from the time value—mostly from variable cost savings in new returns—captured by the value decay parameter. It should also be clear from Figure 9 and (13) that if return policies become even more lenient, i.e., both return rate and percentage new returns increase, and clockspeed continues to increase as well, preponement solutions involving close collaboration with channel partners may become imperative to maintain profitability in small margin businesses. In other words, many of today’s centralized returns handling networks may have to be reengineered in the future.
8. Conclusion

We begin this paper by stating that almost all reverse logistics networks today are driven by efficiency: centralized and focused on economies of scale (local cost minimization through bulk transportation, batching in remanufacturing, large central facilities focused on high utilization, and the like). This paper shows that there are an increasing number of cases where a centralized efficiency-driven reverse network is no longer appropriate. Companies should reconsider the structure of their network, especially if they face large and increasing return rates and high recoverable product value. Return rate and recoverable product value are scale effects, i.e., they impact the magnitude of the costs of the reverse network, and therefore the profitability of the business.

The major parameters driving reverse network design are the time-value decay \( \varphi \) and the proportion of new returns \( p \). Centralized efficient reverse networks are appropriate when both \( \varphi \) and \( p \) are relatively low; that is, when the proportion of unused returned products is low and the product price is relatively stable over time. If we increase the rate at which products lose value over time, it becomes more interesting to consider responsive decentralized return networks and to further increase responsiveness by speeding up transportation, increasing surge capacity at facilities and reducing batching both in transportation and remanufacturing. Saving time will save value, and at some point, will compensate for the losses in economies of scale.

The rate of new returns \( p \) acts as a moderator in that higher values of \( p \) reduce the value of responsiveness in the part of the network between the evaluating facility and the secondary market. On the other hand, higher values of the proportion of new returns \( p \) increase the attractiveness of preponement, because these larger quantities of unused products can then be returned to the forward supply chain faster. Preponement retains the product value of new returns and avoids unnecessary transportation to an evaluating facility before reintroduction in the forward supply chain. The time-value decay parameter \( \varphi \) acts as an amplifier because the speed of reintroduction in the forward supply chain is more critical for products that lose value very quickly. The above insights are qualitatively illustrated in Figure 10.

Our analytic model is quite robust; our simulation built to test the model’s robustness shows that the effects are amplified when capacitated facilities or batching are introduced. Batching tends to slow down the process, and therefore causes greater loss of value when time decay is considerable. Note also that while our model allows for macromanagerial design...
insights, it can also be used for more detailed scenario analyses for a particular company, as we have shown for the HP and Bosch cases.

The major design drivers $p$ and $\varphi$ are influenced by sales conditions (i.e., how liberal the return policies are) and by technological progress (i.e., how fast the technology is changing), respectively. Companies may not be able to influence these parameters to a great extent. Both parameters may increase over time as the speed of technological evolution and the intensity of market competition become more critical for many sectors and products in our global economy.

The implications for management are clear: companies with high return rates and considerable recoverable value should seriously consider redesigning their return networks from a focus on centralization and efficiency to a focus on responsiveness (speed, decentralization) when the rate at which their products lose value is high. If, in addition, many returned products are unused, they should also consider preponement.

References


Blackburn, J. D. 2001. Limits of time-based competition: Strategic sourcing decisions in make-to-stock manufacturing. Working paper 01-19, Owen Graduate School of Management, Vanderbilt University, Nashville, TN.


