Taking market forces into account in the design of production-distribution networks: A positioning by anticipation approach

Didier Vila

Alain Martel

and

Robert Beauregard

Working Paper DT-2005-AM-4

Accepted for publication in

The Journal of Industrial and Management Optimization

Research Consortium on e-Business in the Forest Products Industry (FORAC),

Network Organization Technology Research Center (CENTOR),

Université Laval, Québec, Canada

Taking market forces into account in the design of production-distribution networks: A positioning by anticipation approach

Didier Vila [1,2], *Alain Martel* [1 *] *and Robert Beauregard* [1]

(1) Université Laval, FOR@C Research Consortium, Network Organization Technology Research Center (CENTOR), Sainte-Foy, Québec, G1K7P4, Canada.
(2) École Nationale Supérieure des Mines de Saint-Étienne, Centre G2I, 158 cours Fauriel, 42023 Saint-Étienne cedex 2, France.

Abstract. This paper presents an approach to take into account market opportunities when designing production-distribution networks. Three types of sub-markets found in several industrial contexts are analyzed: spot markets, contracts and Vendor Managed Inventory (VMI) agreements. For contracts and VMI agreements, customer preferences with respect to different logistics policies are considered. A price-supply function is proposed to model the spot market behavior. The production-distribution network design problem is formulated as a two-stage stochastic program with fixed recourse. Finally, a sample average approximation method (SAA), based on Monte Carlo sampling techniques, is used to solve the model.

Keywords. Production-distribution Network Design, Mathematical Programming, Monte-Carlo Sampling Methods, Market Analysis, Logistics Policy Selection.

Acknowledgements. This project would not have been possible without the financial support and the collaboration of FOR@C's partners. The information on discrete choice models provided by D. Bolduc (Economics Department, Université Laval) was also very helpful.

* Corresponding author E-mail: Alain.Martel@centor.ulaval.ca

1 Introduction

The performance of a supply chain for a given product-market depends critically on the structure of its production-distribution network, i.e. the number, location, mission, technology and capacity of the facilities of the firms involved, but also on its capacity to make winning offers to its potential customers. A supply chain structure leading to lower prices, better service and better quality products than those of competitors leads to higher market shares and thus to higher revenues. By assuming that the demand for products is predetermined, classical network design models overlook this important aspect of the problem. The exact nature of the network design problems encountered in practice depends very much on the industrial context in which they occur, and on the breath of the markets considered. Networks covering several countries lead to much more complex design problems because factors such as exchange rates, duties and income tax must be taken into account. This paper presents a generic methodology to explicitly consider market forces when designing international production-distribution networks for make-to-stock products.

Logistics network design problems integrate location, capacity acquisition and technology selection sub-problems. A review of the initial literature on these problems is found in Verter and Dincer (1992). The first location-allocation model proposed (Geoffrion and Graves, 1974) was a single echelon single period model to determine the distribution centers to use, as well as the assignment of products and clients to these centers, in order to minimise the total cost of the system in a domestic context. Several extensions to this model were subsequently made to take into account multiple echelons (Cohen and Lee, 1989; Pirkul and Jayaraman, 1996; Martel and Vankatadri, 1999; Vidal and Goetschalckx, 2001), multiple production seasons (Cohen *et al*., 1989; Arntzen *et al*., 1995; Dogan and Goetschalckx, 1999), capacity acquisition and technology selection (Eppen *et al*., 1989; Verter and Dincer, 1995; Mazzola and Neebe, 1999; Paquet *et al*., 2004; Martel, 2005), economies of scale (Cohen and Moon, 1990, 1991; Mazzola and Schantz, 1997; Martel and Vankatadri, 1999), after tax net revenue maximization in an international context (Cohen *et al*., 1989; Arntzen *et al.,* 1995; Vidal and Goetschalckx, 2001) and product development and recycling (Fandel and Stammen, 2004). Geoffrion and Powers (1995) and Shapiro *et al*. (1993) discuss the evolution of strategic supply chain design models and Vidal and Goetschalckx (1997) present many of these models. A modeling framework integrating most of these results is presented in Martel (2005).

In most industrial sectors, the market is not monolithic and several product-markets governed by different rules-of-the-game can be found. For example, several natural resource based products, such as lumber, can be sold on the spot market or through contracts with major customers. In the later case, the probability of getting a contract depends on a set of qualifying and order-winning criteria such as price, lead-time and fill-rate. For a given potential customer, a company is able to win on several of these criteria only if its production-distribution facilities are better positioned than those of its competitors. Despite the obvious impact production-distribution network structures can have on company performance in such contexts, little work has been done to take market forces into account explicitly in network design models. Shapiro (2001) stresses the necessity to integrate strategic marketing and production-distribution decisions in the same model to design superior supply chains. In their literature review on the modeling of supply chain contracts, Tsay *et al.* (1999) do not include any papers dealing with production-distribution networks design issues. Rosenfield *et al*. (1985) show how the performance of different logistic network designs can be characterized by an efficient cost-service frontier. Starting from these results, this paper develops a generic approach and a two-stage stochastic programming model to design production-distribution networks improving the competitive position of a company on its markets. More specifically, three types of sub-markets found in several industrial contexts are considered: spot markets, contracts and Vendor Managed Inventory (VMI) agreements.

The rest of the paper is organized as follows. In the next section, the design approach proposed is explained and the stochastic programming model on which it is based is formulated. The following section presents the sample average approximation model, based on Monte Carlo sampling techniques, proposed to obtain network designs. Finally, numerical experiment results to test the approach are presented and analyzed.

2 Methodology

Without loss of generality, to simplify the presentation, we restrict ourselves to the case of a single echelon production-distribution network of the type illustrated in Figure 1. It is assumed that the production-distribution sites $s \in S^{pd}$ are already in use and that a subset of potential distribution sites $s \in S^{d}$ must be selected. In order to address the problem considered, it is necessary to understand the chronology of events underlying the design process. As illustrated in Figure 2, the hierarchical planning and execution process proposed involves four steps which are explained in the next subsections.

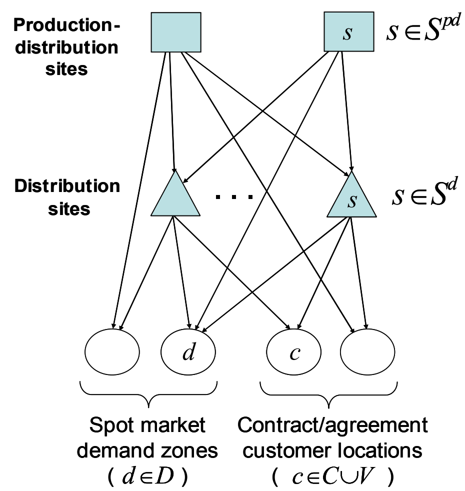

Figure 1: Network Structure

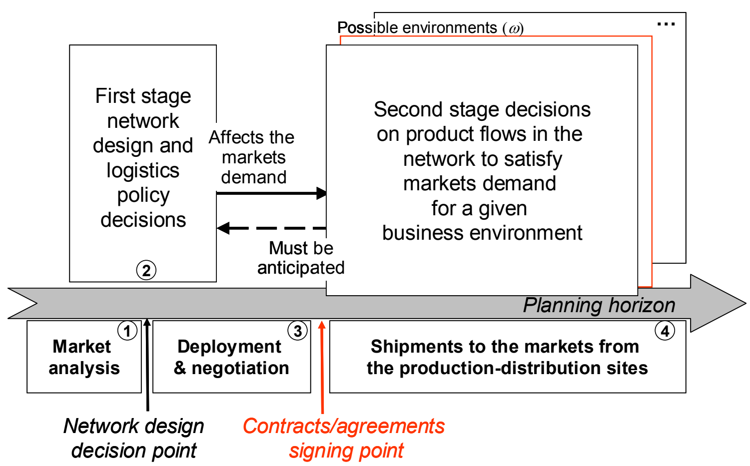

Figure 2: Chronology of Events

2.1 Market segmentation and logistics policies definition

It is assumed that the company is selling products $p \in P$ in several countries $o \in O$ and that each national division covers a set of distinct product-markets $m \in M_o$, $o \in O$. The set of national product-markets M_o can be partitioned into three sub-sets:

- A set of spot markets $m \in SM_o$. A spot market m is characterized by products $p \in P_m \subset P$, by demand zones $d \in D_{pm}$ and by decreasing price step functions $P_{pm}(x_{pm})$ of the total sales x_{pm} of product *p* in market *m* (see Figure 3). A demand zone d is a geographical aggregate of several ship-to locations with coordinates associated to its centroid. We use $m(d)$ to identify the spot market *m* to which demand zone *d* belongs.

- A set of customer contracts $c \in C_o$, which is partitioned into a set PC_o of potential contracts and a set SC_o of signed contracts. A contract *c* is an engagement to deliver a predetermined quantity x_c of product $p_c \in P$ to a given ship-to location, during a predetermined period of time, and with guaranteed price and lead time.
- A set of Vendor Managed Inventory (VMI) agreements $c \in V_o$ which is partitioned into a set PV_o of potential VMI agreements and a set SV_o of signed VMI agreements. A VMI agreement *c* is an engagement to deliver a predetermined quantity x_c of product $p_c \in P$ to a given ship-to location, during a predetermined period of time (assumed to be the same for all contracts and agreements), and with guaranteed price and fill rate.

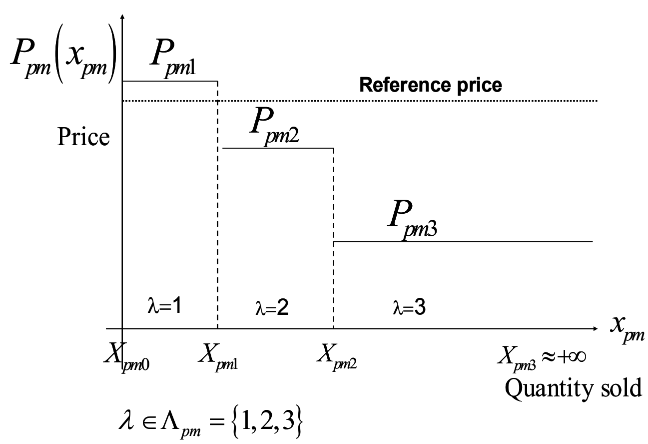

Figure 3: Spot Market Price Step Function

The spot market can be considered as a *recourse* which can absorb any amount of product, but for a price decreasing with quantity. Signed contracts/agreements yield a deterministic demand to be satisfied, but potential contracts/agreements define a stochastic demand process. Additional flexibility is also possible through product substitution: indeed, in all markets, a product $p \in P$ can be substituted by a product $p' \in SP^p$, SP^p being a set of substitutes for product p.

The decreasing price step function associated to product $p \in P_m$ on spot market $m \in SM_o$ has the form:

$$P_{pm}\left(x_{pm}\right) = P_{pm\lambda} \quad \text{if} \quad X_{pm\lambda-1} < x_{pm} \le X_{pm\lambda}, \quad \lambda \in \Lambda_{pm} \quad \text{with} \quad X_{pm0} = 0, X_{pm\lambda_{pm}} = +\infty, \lambda_{pm} \triangleq \left|\Lambda_{pm}\right|$$

where Λ_{pm} is the set of steps of the function, $P_{pm\lambda}$ is the unit price for step $\lambda \in \Lambda_{pm}$ and $X_{pm\lambda}$ is the upper bound of step $\lambda \in \Lambda_{pm}$, as illustrated in Figure 3. We assume that prices and step lengths are determined by the company using price forecasts based on the historical behavior of the firm prices on the spot market, in relation with an expected reference price.

Because of competition, it is assumed that potential customers will sign contracts only if the company can demonstrate that it has the resources required to comply with all the clauses of the contracts/agreements. Consequently, the production-distribution network must be designed to satisfy signed contracts and agreements $c \in SC_o \cup SV_o$ and to be in a position to satisfy some potential new contracts and agreements $c \in PC_o \cup PV_o$, knowing that the uncommitted production can be sold on the spot market.

In order to win contracts/agreements, the company has to develop different offers to satisfy potential customers better than its competitors. Following Hill (1994), it is assumed that these offers must be defined in terms of criteria that win contracts on the marketplace (order winners) and criteria that qualify the company as a potential supplier (qualifiers). These offers are formalized here through the concept of *logistics policy*. A logistic policy i is characterized by a vector of target values for the relevant order winning and qualifying criteria and by the fix marketing and logistics cost K_i incurred when the policy is implemented. Let i_c be the logistics policy implemented for signed contract/agreement c and I_c be the set of policies considered for potential customer location c. For convenience, we define $I_c = \{i_c\}$, $c \in SC_o \cup SV_o$ and $I = \cup_c I_c$, and we use $c(i)$ to identify the customer location c to which policy i applies. Without loss of generality, is assumed in this text that a policy $i \in I_c$ considered for customer contract c is characterized by two order winning criteria, namely the product price P_i and the maximum delivery time υ_i. Similarly, it is assumed that the order winners associated to a policy $i \in I_c$ considered for VMI agreement *c* are the product price P_i and the fill rate α_i. The fill rate α_i relates to the necessity to keep a safety stock at the customer location. The inventory holding cost of this safety stock for the contract period is included in the fix policy cost K_i.

For a given logistics policy, it may not be possible to satisfy the target values specified for the order winners and the qualifiers from all the production and distribution facilities in the network. For example, for a policy $i \in I_c$, $c \in PC_o$, if the delivery time required to service customer location $c(i)$ from a facility $s \in S^{pd} \cup S^d$ is longer than υ_i, then this facility cannot be used to implement the policy. This leads to the association of a set of *admissible sites* $S_i^i \subseteq S^{pd} \cup S^d$ to each logistics policy *i*. S_i^i, $i \in I_c$, is the set of facilities $s \in S^{pd} \cup S^d$ the company could use to comply with the terms of logistics policy *i*. Note that the selection of a logistics policy $i \in I_c$ does not imply that the potential contract or VMI agreement *c* will be signed, but it qualifies the company to bid for this contract. The probability θ_i that contract or VMI agreement $c \in PC_o \cup PV_o$ will be signed if logistics policy $i \in I_c$ is selected can however be evaluated.

Discrete choice analysis can be used to estimate the probabilities θ_i, $i \in I$, using econometric models (Ben-Akiva and Lerman, 1985). This approach is based on the modeling of customer preferences among a set of offers which could be made by the company or by its competitors. Each offer corresponds to a list of order winning criteria target values, coupled with the identity of the company making the offer. For contracts, an offer *i* would thus correspond to the triplet $(P_i, \upsilon_i, id(i))$, where *id*(*i*) is the identity of the company making the offer. Let the set of offers to the customer associated to potential contract $c \in PC_o$ be denoted by $Offer_c$. Note that this set necessarily includes all the logistics policies considered for potential contract *c*, i.e. $Offer_c \supset I_c$. Based on random utility theory, the utility $U_c(i)$ of an offer $i \in Offer_c$ for customer $c \in PC_o$ can be modeled with the linear function:

$$U_c(i) = \beta_{socio} o(c) + \beta_{price} P_i + \beta_{delay} \upsilon_i + \beta_{comp} id(i) + \varepsilon_{ic}$$

where β_{socio}, β_{price}, β_{delay} and β_{comp} are parameters to be estimated and where ε_{ic} is a random component. The independent random variables ε_{ic} are Gumbel-distributed with a location parameter η and a scale parameter $0 < \mu$, i.e. they have the same probability density function $f(\varepsilon) = \mu e^{-\mu(\varepsilon-\eta)} \exp(-e^{-\mu(\varepsilon-\eta)})$. The parameter β_{socio} associated to the country $o(c)$ of customer c, captures local socio-economic effects. In order to estimate the model parameters, the *stated* preferences framework proposed by Louviere *et al*. (2000) can be used. A questionnaire with hypothetical offers is submitted to a sample of customers. With these observations, maximum likelihood estimators are used to obtain the parameters value. This can be implemented, for example, with the BIOGEME software developed by Bierlaire and available on the Web at http://roso.epfl.ch/biogeme.

Neoclassical utility theory is based on the premise that decision-makers chose their highest utility options. In our context, this leads to the assumption that the probability that the customer associated to potential contract $c \in PC_o$ would choose an offer $i \in Offer_c$ is given by:

$$P(i \mid Offer_c) = P(U_c(i) \geq U_c(i'), \forall i' \in Offer_c)$$

When using a Multinomial Logit discrete choice model, this probability can be calculated with the expression:

$$P(i \mid Offer_c) = \frac{e^{\mu(\beta_{socio} o(c) + \beta_{price} P_i + \beta_{delay} \upsilon_i + \beta_{comp} id(i))}}{\sum_{i' \in Offer_c} e^{\mu(\beta_{socio} o(c) + \beta_{price} P_{i'} + \beta_{delay} \upsilon_{i'} + \beta_{comp} id(i'))}}$$

However, since a single logistic policy $i \in I_c \subset Offer_c$ could eventually be offered to potential customer c, in order to calculate the probability θ_i, the only offers to consider are offer i and the offers of competitors. Let $Offer_c(i)$ be the subset of $Offer_c$ constituted of the competitor offers and of offer $i \in I_c \subset Offer_c$. Then, the probability that contract $c(i)$ would be signed, if logistics policy $i \in I_c$ is selected, is given by:

$$\theta_i = P(i \mid Offer_{c(i)}(i)) = \frac{e^{\mu(\beta_{socio} o(c(i)) + \beta_{price} P_i + \beta_{delay} \upsilon_i + \beta_{comp} id(i))}}{\sum_{i' \in Offer_{c(i)}(i)} e^{\mu(\beta_{socio} o(c(i)) + \beta_{price} P_{i'} + \beta_{delay} \upsilon_{i'} + \beta_{comp} id(i'))}}$$

The same approach can be used to obtain the probabilities θ_i that VMI agreements will be signed. It is through this probability estimation procedure that competitor potential actions are taken into account in our production-distribution network design methodology.

2.2 Network design decisions and anticipated shipping decisions

The goal of the company is to design its production-distribution network anticipating the future by simultaneously selecting adequate logistics policies, and by allocating production capacity and locating distribution centers to support these policies. This requires the definition of the following decision variables:

X_{ps} = Quantity of product p produced in production-distribution center $s \in S^{pd}$.

Y_s = Binary variable equal to 1 if potential distribution center $s \in S^d$ is used and to 0 otherwise.

Z_i = Binary variable equal to 1 if logistics policy $i \in I_c$, $c \in PC_o \cup PV_o$ is selected and to 0 otherwise.

For convenience, we also define the following design variable vectors:

$$\mathbf{X} \triangleq \left[X_{ps} \right], \ \mathbf{Y} \triangleq \left[Y_s \right] \text{and } \mathbf{Z} \triangleq \left[Z_i \right].$$

In order to design the network properly, the impact of these decisions on future market demand and on the operational costs associated to the delivery of products sold to customer locations must be anticipated. In order to anticipate future network costs and revenues, we assume that, as illustrated in Figure 2, once design decisions have been implemented, at some point in time, customers will accept or reject the company’s offers and the quantity of products to ship in the contract/agreement period will be known. A particular reaction from the customers to the company’s potential offers defines a business *environment*. Although, it is clear that all the contracts/agreements would not be signed or rejected at the same point in time, we propose to anticipate the impact of the design by computing expected network flow costs and revenues during a predetermined contract duration period for all the future environments the company could face.

Let Ω be the set of all possible environments. An *environment* $\omega \in \Omega$ is a binary variable vector of dimension $|I|$ indicating whether the customers would sign a contract/agreement ($\omega_i = 1$) or not ($\omega_i = 0$) for all possible logistics policies $i \in I$. Note that $\omega_i = 1$, $c(i) \in SC_o \cup SV_o$, in all environments since these contracts/agreements are already signed when the design decisions are made. Also note that it is not necessary to include the spot market explicitly in the description of an environment since it is considered as a recourse which can absorb any outstanding production.

Since ω_i is a binary variable, the number of possible environments which could be observed is given by 2^n, $n = \Sigma_{o \in O}(\Sigma_{c \in PC_o \cup PV_o} |I_c|)$. Since the company cannot implement more than one logistics policy $i \in I_c$ at the same time, for a given environment $\omega \in \Omega$, the demand on the contract and VMI agreement markets in country $o \in O$ is given by:

$$(\sum_{i \in I_c} \omega_i Z_i) x_c \quad c \in PC_o \cup PV_o; \quad x_c \quad c \in SC_o \cup SV_o$$

Also, the probability $p(\omega)$ that a given environment $\omega \in \Omega$ will prevail, can be derived from the probabilities θ_i of signing a contract or VMI agreement under logistics policies $i \in I$. This clearly shows that market demand depends on logistic network design decisions. In other words, in our approach, logistic network design decisions are not made simply to adapt to a predetermined demand, they are strategic competitive positioning decisions used to influence customer behavior.

The network design decisions **X**, **Y** and **Z** considered here are *first stage* decisions which would normally be implemented in practice immediately after they are made. Their implementation would not be instantaneous however. In particular, the decisions Y lead to the redeployment of the company distribution network and the decisions Z to the negotiation of contracts with potential customers, which may take several months. As illustrated in Figure 2, at the end of this implementation phase, the environment $\omega \in \Omega$ in which the logistics network implemented will be used is revealed and shipment decisions to satisfy market demand during the contracts/agreements period must be made. Although these *second stage* decisions would not be implemented *per se* in practice, they are important to anticipate the impact of the network design on network flow costs and revenues. These second stage decisions are clearly dependent on the environment $\omega \in \Omega$ which will eventually prevail. Taking product substitution possibilities into account, this leads to the definition of the following network flow decision variables:

$F_{pss'}(\omega)$ = Flow of product p between production-distribution site $s \in S^{pd}$ and distribution site $s' \in S^{d}$ for environment $\omega \in \Omega$.

$F_{pp'sd\lambda}(\omega)$ = Outbound flow from site s of product $p' \in P$ used to satisfy the demand for product $p \in P$ in demand zone d of spot market $m(d)$, and sold at price $P_{pm(d)\lambda}$ of step function interval $\lambda \in \Lambda_{pm(d)}$, for environment $\omega \in \Omega$.

$F_{psi}(\omega)$ = Outbound flow from site s of product $p \in P$ used to satisfy the demand of product $p_{c(i)}$ at customer location $c(i) \in C_o \cup V_o$ with the logistics policy i, for environment $\omega \in \Omega$.

For convenience, we define the vector of second stage variables

$$\mathbf{F}(\omega) \triangleq \left[F_{pss'}(\omega), F_{pp'sd\lambda}(\omega), F_{psi}(\omega) \right].$$

The design approach described in the previous paragraphs leads to the formulation of the problem as a two-stage stochastic program with recourse (Birge and Louveaux., 1997).

2.3 Stochastic programming model formulation

In order to formulate the demand, distribution and manufacturing constraints of the model, the following sets are required:

SP_p = Set of products which product p can substitute for.

S^o_{ps} = Set of distribution sites (output destinations) which can receive product *p* from production-distribution site *s*.

S^i_{ps} = Set of production-distribution sites (input sources) which can ship product *p* to distribution site *s*.

S^i_m = Set of facilities which can ship products $p \in P_m$ to spot market *m* ($S^i_m \subseteq S^d \cup S^{pd}$).

SM_{ps} = Set of spot markets which can receive substitute products *p* from node *s*, i.e. $SM_{ps} = \left\{ m \middle| s \in S^i_m, p \in \cup_{p' \in P_m} SP_{p'} \right\}$.

For an environment $\omega \in \Omega$, the flow of products or substitute products from the production and distribution sites must cover the demand associated to the signed contracts and VMI agreements. Knowing that logistics policy i_c is used for customer $c \in SC_o \cup SV_o$, this gives rise to the following constraints:

$$\sum_{p \in SP^{Pc}} \sum_{s \in S^i_{i_c}} F_{psi_c}(\omega) = x_c \quad \omega \in \Omega,\ c \in \cup_{o \in O} SC_o \cup SV_o \qquad 1)$$

Concerning potential customers $c \in PC_o \cup PV_o$, the form taken by the demand constraints depends on the customer response for the environment $\omega \in \Omega$ considered and on first stage logistics policy implementation decisions Z_i. The constraints required are the following:

$$\sum_{p \in SP^{Pc}} \sum_{s \in S^i_i} F_{psi}(\omega) = x_c Z_i \quad \omega \in \Omega,\ c \in \cup_{o \in O} PC_0 \cup PV_0,\ i \in I_c(\omega) \qquad 2)$$

$$\sum_{i \in I_c} Z_i \leq 1 \quad c \in \cup_{o \in O} PC_o \cup PV_o \qquad 3)$$

where $I_c(\omega) = \{ i \in I_c | \omega_i = 1 \}$. In 2), for each environment $\omega \in \Omega$, the demand constraints are included only for the logistics policies $i \in I_c(\omega)$ which would lead to a signature of the contract/agreement. For the other policies, the contract would not be signed and hence the demand would be zero. This could be included as explicit constraints but it is more efficient to drop these constraints and the associated flow variables. Also note that because there is a fixed cost K_i associated to Z_i in the objective function, in the optimal solution $Z_i = 0,\ i \notin \cup_{\omega \in \Omega} I_c(\omega)$.

Constraints 3) are required to make sure that at most one logistics policy i in I_c will be implemented.

As indicated earlier, a spot market m is composed of demand zones $d \in D_{pm}$ for each product $p \in P_m$. We assume that prices on a spot market m are based on the sales volume x_{pm} of product $p \in P_m$ on that market, but that the proportions κ_d, $d \in D_{pm}$, of the sales x_{pm} made in each demand zone $d \in D_{pm}$ are known. Given that, for an environment, upper bounds on the flow of products to demand zone $d \in D_{pm}$, for each step of the price function, are given by:

$$\sum_{p' \in SP^p} \sum_{s \in S_m^i} F_{pp'sd\lambda}(\omega) \le \kappa_d (X_{pm\lambda} - X_{pm(\lambda-1)}) \quad \omega \in \Omega, m \in \cup_{o \in O} SM_o, p \in P_m, d \in D_{pm}, \lambda \in \Lambda_{pm} \qquad 4)$$

Since we want to maximize profits, with the type of price step functions used, it can be shown (see the Appendix for a proof) that any *optimal solution* to the model will be such that, for step $\lambda \in \Lambda_{pm}$:

$$\sum_{p' \in SP^p} \sum_{s \in S_m^i} F_{pp'sd\lambda}(\omega) > 0 \Rightarrow \sum_{p' \in SP^p} \sum_{s \in S_m^i} F_{pp'sd\lambda'}(\omega) = \kappa_d (X_{pm\lambda'} - X_{pm(\lambda'-1)}), \quad \lambda' < \lambda$$

For this reason, it is sufficient to include constraints 4) in the model to ensure that spot market prices will be formulated properly.

The facilities capacity can be modeled using the following parameters:

b_{ps} = Quantity of product p which can be produced in production center s.

b_s = Warehousing capacity of distribution center s in an appropriate unit.

q_p = Warehousing capacity consumption rate per unit of product p.

Using this notation, the capacity constraints of the production-distribution and distribution facilities are formulated as follows:

$$X_{ps} \le b_{ps} \quad p \in P, s \in S^{pd} \qquad 5)$$

$$\sum_{p \in P} q_p \sum_{s' \in S_{ps}^i} F_{ps's}(\omega) \le b_s Y_s \quad \omega \in \Omega, s \in S^d \qquad 6)$$

For each environment, the following flow conservation constraints must also hold:

$$\sum_{s' \in S_{ps}^{o}} F_{pss'}(\omega) + F_{ps}(\omega) = X_{ps} \quad \omega \in \Omega,\ p \in P,\ s \in S^{pd} \tag{7)}$$

$$F_{ps}(\omega) = \sum_{s' \in S_{ps}^{i}} F_{ps's}(\omega) \quad \omega \in \Omega,\ p \in P,\ s \in S^{d} \tag{8)}$$

$$F_{ps}(\omega) = \sum_{m \in SM_{ps}} \sum_{p' \in P_m} \sum_{d \in D_{pm}} \sum_{\lambda \in \Lambda_{p'm}} F_{p'psd\lambda}(\omega) + \sum_{c \in C \cup V | p \in SP_{pc}} \sum_{i \in I_c(w) | s \in S_i^{i}} F_{psi}(\omega) \quad \omega \in \Omega,\ p \in P,\ s \in S^{pd} \cup S^{d} \tag{9)}$$

where $F_{ps}(\omega)$ is a working decision variable defined by 9) and used to simplify the formulation.

To calculate the total revenues and costs of a network design, the following financial parameters are required:

A_s = Fixed cost of using distribution site $s \in S^d$ for the planning horizon.

c_{ps} = Cost of producing one unit of product p on site $s \in S^{pd}$.

f_{psn}^{o} = Unit cost of the flow of product p between site s and node n paid by origin s (this cost includes the customer-order processing cost, the shipping cost, the variable transportation cost and the inventory-in-transit holding cost).

f_{psn}^{t} = Unit transportation cost of product p from site s to node n (this cost is included in f_{psn}^{o}).

f_{pns}^{d} = Unit cost of the flow of product p between node n and site s paid by destination s (this cost includes the supply-order processing costs and the receiving cost).

h_{ps} = Unit inventory holding cost of product p in facility s.

π_{ps} = Transfer price of product p shipped from site s.

$e_{oo'}$ = Exchange rate, i.e. number of units of country o currency by units of country o' currency (the index $o = 0$ is given to the base currency).

δ_{pns} = Import duty rate applied to the CIF price of product p when transferred from the country of node n to the country of site s.

τ_o = Income tax rate of country *o*.

ρ_{ps} = Number of seasons of inventory (order cycle and safety stocks) of product p held at site s.

In an international context, in order to take transfer prices and taxes into account correctly, it is necessary to derive an income statement for each network facility. Let:

$C_s(\omega)$ = Total site s expenses for the planning horizon under environment $\omega \in \Omega$.

$R_s(\omega)$ = Total site s revenues for the planning horizon under environment $\omega \in \Omega$.

Then, using the expenditure and revenue elements in Table 1, where $o(s)$ denotes the country of site s, it is seen that:

$$C_s(\omega) = c) + e) + f) + g) \quad s \in S^{pd}, \quad \omega \in \Omega \qquad 10)$$

$$C_s(\omega) = a) + b) + d) + e) + g) \quad s \in S^{d}, \quad \omega \in \Omega \qquad 11)$$

$$R_s(\omega) = h) + i) \quad s \in S^{pd}, \quad \omega \in \Omega \qquad 12)$$

$$R_s(w) = i) \quad s \in S^{d}, \quad \omega \in \Omega \qquad 13)$$

The operating income for each national division $o \in O$, under a given environment ω, taking into account the fixed costs of all the logistics policies considered for contracts and VMI agreements $c \in C_o \cup V_o$, is given by:

$$M_o(\omega) = \sum_{s \in S_o} (R_s(\omega) - C_s(\omega)) - \sum_{c \in C_o \cup V_o} \sum_{i \in I_c} K_i Z_i \qquad 14)$$

where S_o is the set of sites located in country *o*. We must distinguish positive margins from negative margins because there is no income tax to pay on losses. To do this, $M_o(\omega)$ must be separated in its negative and positive parts by defining the operating income $M_o(\omega) = M_o^+(\omega) - M_o^-(\omega)$. Given this, the objective function of the model, i.e. the after tax net revenue of the corporation in its reference currency, is given by the expression $\sum_{0 \in O} e_{0o}[(1-\tau_0)M_o^+(\omega) - M_o^-(\omega)]$.

<table>
<tr><th></th><th></th><th>Distribution center
($s \in S^d$)</th><th>Production-distribution center
($s \in S^{pd}$)</th></tr>
<tr><td rowspan="7">Expenses</td><td>a) Inflow transfer cost</td><td>$\sum_{p\in P}\sum_{s'\in S^i_{ps}}(1+\delta_{pss'})e_{o(s)o(s')}(\pi_{ps'}+f^t_{ps's})F_{ps's}(\omega)$</td><td></td></tr>
<tr><td>b) Receptions from other sites</td><td>$\sum_{p\in P}\sum_{s'\in S^i_{ps}} f^d_{ps's}F_{ps's}(\omega)$</td><td></td></tr>
<tr><td>c) Production</td><td></td><td>$\sum_{p\in P} c_{ps}X_{ps}$</td></tr>
<tr><td>d) Facilities options cost</td><td>A_sY_s</td><td></td></tr>
<tr><td>e) Order cycle and safety stocks</td><td></td><td>$\sum_{p\in P} h_{ps}\rho_{ps}(F_{ps}(\omega)+\sum_{s'\in S^o_{ps}} F_{pss'}(\omega))$</td></tr>
<tr><td>f) Outflows to other sites</td><td colspan="2">$\sum_{p\in P}\sum_{s'\in S^o_{ps}} f^o_{pss'}F_{pss'}(\omega)$</td></tr>
<tr><td>g) Outflows to demand zones</td><td colspan="2">$\sum_{c\in C\cup V}\sum_{i\in I_c(\omega) \mid s\in S^i_i}\sum_{p'\in SP^{p_c}} f^o_{p'sd_c}F_{p'si}(\omega)+$
$\sum_{m\in SM \mid s\in S^i_m}\sum_{p\in P_m}\sum_{d\in D_{pm}}\sum_{\lambda\in\Lambda_{pm}}\sum_{p'\in SP^p} f^o_{p'sd}F_{pp'sd\lambda}(\omega)$</td></tr>
<tr><td rowspan="2">Revenues</td><td>h) Outflows to other sites</td><td></td><td>$\sum_{p\in P}\sum_{s'\in S^o_{ps}}(\pi_{ps}+f^t_{pss'})F_{pss'}(\omega)$</td></tr>
<tr><td>i) Outflows to demand zones</td><td colspan="2">$\sum_{c\in C\cup V} e_{o(s),o(d_c)}\sum_{i\in I_c(\omega) \mid s\in S^i_i}\sum_{p'\in SP^{p_c}} P_iF_{p'si}(\omega)$
$+\sum_{m\in SM \mid s\in S^i_m}\sum_{p\in P_m}\sum_{d\in D_{pm}} e_{o(s),o(d)}\sum_{\lambda\in\Lambda_{pm}}\sum_{p'\in SP^p} P_{pm\lambda}F_{pp'sd\lambda}(\omega)$</td></tr>
</table>

Table 1: Facilities Expenses and Revenues in Local Currency for a Given Environment

Based on the previous discussion, it can be seen that the stochastic program to solve is the following:

$$\text{Max } \sum_{\omega\in\Omega} p(\omega)\sum_{0\in O} e_{0o}[(1-\tau_0)M^+_o(\omega)-M^-_o(\omega)] \quad \text{(O-1)}$$

subject to constraints 1) to 13), to the national divisions operating income definitions

$$\sum_{s \in S_o} (R_s(\omega) - C_s(\omega)) - \sum_{c \in C_o \cup V_o} \sum_{i \in I_c} K_i Z_i - M_o^+(\omega) + M_o^-(\omega) = 0, \quad o \in O,\ \omega \in \Omega, \qquad 15)$$

and to the non-negativity constraints:

$$X_{ps} \geq 0\ p \in P, s \in S^{pd} \quad Y_s \in \{0,1\}\ s \in S^d \quad Z_i \in \{0,1\}\ o \in O, c \in PC_o \cup PV_o, i \in I_c$$

$$F_{pss'}(\omega) \geq 0\ p \in P, s \in S^{pd}, s' \in S_{ps}^o, \omega \in \Omega \quad F_{ps}(\omega) \geq 0\ p \in P, s \in S^{pd} \cup S^d, \omega \in \Omega$$

$$F_{p'si}(\omega) \geq 0\ o \in O, c \in C_o \cup V_o, p' \in SP^{p_c}, \omega \in \Omega, i \in I_c(\omega), s \in S_i^i$$

$$F_{pp'sd\lambda}(\omega) \geq 0\ o \in O, m \in MS_o, p \in P_m, p' \in SP^p, s \in S_m^i, d \in D_{pm}, \lambda \in \Lambda_{pm}, \omega \in \Omega$$

$$C_s(\omega) \geq 0, s \in S^{pd} \cup S^d,\ \omega \in \Omega \quad R_s(\omega) \geq 0, s \in S^{pd} \cup S^d,\ \omega \in \Omega$$

$$M_o^+(\omega) \geq 0, o \in O,\ \omega \in \Omega \quad M_o^-(\omega) \geq 0, o \in O,\ \omega \in \Omega$$

Note that in some contexts, companies may prefer to maximize corporate net revenues *before* taxes in the reference currency, that is $\sum_{0 \in O} e_{0o} M_o(\omega)$. When this is the case, constraints 15) are not necessary and 14) can be substituted back into the net revenue expression, to get the following objective function:

$$\text{Max} \sum_{0 \in O} e_{0o} \{ \sum_{s \in S_o} \sum_{\omega \in \Omega} p(\omega)[R_s(\omega) - C_s(\omega)] - \sum_{c \in C_o \cup V_o} \sum_{i \in I_c} K_i Z_i \} \qquad \text{(O-2)}$$

Furthermore, the revenue $R_s(\omega)$ and expenditure $C_s(\omega)$ definitions 10) to 13) can also be substituted back into the objective function, which decreases the size of the model significantly.

3 Stochastic Programming Model Solution

3.1 Solution approach

Under objective function (O-2), the two-stage stochastic programming model with fixed recourse formulated to design logistics networks has the following form:

$$\begin{aligned} &\max f(\mathbf{X},\mathbf{Y},\mathbf{Z}) = \textstyle\sum_{\omega \in \Omega} p(\omega)\left[\max \mathbf{qF}(\omega)\right] - \left(\mathbf{cX} + \mathbf{aY} + \mathbf{kZ}\right) \qquad \text{SP} \\ &s.t.\ \ \mathbf{A}\left[\mathbf{X},\mathbf{Z}\right] \leq \mathbf{b}, \\ &\quad -\mathbf{T}(\omega)\mathbf{Z} + \mathbf{UF}(\omega) = \mathbf{0}, \quad \omega \in \Omega, \\ &\quad -\mathbf{V}\left[\mathbf{X},\mathbf{Y}\right] + \mathbf{WF}(\omega) = \mathbf{h}, \quad \omega \in \Omega, \\ &\quad \mathbf{X} \geq \mathbf{0}; \mathbf{Y},\mathbf{Z}\ \text{binary}; \mathbf{F}(\omega) \geq \mathbf{0}, \quad \omega \in \Omega. \end{aligned}$$

where $\mathbf{q}$, $\mathbf{c}$, $\mathbf{a}$ and $\mathbf{k}$ are vectors of revenues and costs, $\mathbf{A}$, $\mathbf{T}(\omega)$, $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{W}$ are matrices of parameters, $\mathbf{b}$ and $\mathbf{h}$ are right hand side vectors and $\mathbf{0}$ is a null vector. As it stands, this stochastic program is difficult to solve because the number of possible environments in Ω can be extremely large. Fortunately, there is no second stage binary variable, which means that each environment adds only continuous variables. Nevertheless, for a practical case, program (SP) could include billions of second stage variables, which is prohibitive even for the very efficient solvers currently available. Note that the model obtained with objective (O-1) is similar to SP and that the solution approach proposed in the next paragraphs applies to both cases.

The approach proposed to solve (SP) seeks to find the best possible design with the mathematical programming solvers currently available. The approach used is based on the Monte Carlo sampling methods described by Shapiro (2003). A random sample of environments is generated outside the optimization procedure and then a sample average approximation (SAA) program is constructed and solved. The idea is first to generate an independent identically distributed sample of N environments $\left\{\omega^1,...,\omega^N\right\} = \Omega^N \subset \Omega$ from the initial distribution of ω. Then the SAA program to solve is the following:

$$\begin{aligned}
\max = & \frac{1}{N}\sum\nolimits_{\omega\in\Omega^N} \mathbf{qF}(\omega) - \left(\mathbf{cX}+\mathbf{aY}+\mathbf{kZ}\right) \qquad \text{SAA}(\Omega^N)\\
s.t.\quad & \mathbf{A}\left[\mathbf{X},\mathbf{Z}\right] \le \mathbf{b},\\
& \text{-}\mathbf{T}(\omega)\mathbf{Z}+\mathbf{UF}(\omega)=\mathbf{0}, \quad \omega\in\Omega^N,\\
& \text{-}\mathbf{V}\left[\mathbf{X},\mathbf{Y}\right]+\mathbf{WF}(\omega)=\mathbf{h}, \quad \omega\in\Omega^N,\\
& \mathbf{X}\ge\mathbf{0}; \mathbf{Y},\mathbf{Z}\ \text{binary}; \mathbf{F}(\omega)\ge\mathbf{0}, \quad \omega\omega\in\Omega^N.
\end{aligned}$$

Clearly, the quality of the solution obtained with this approach improves as the size N of the sample of environments used increases. The approach suggested here is to use a sample size N giving a SAA program solvable in a reasonable time with a commercial solver such as CPLEX. The SAA program is solved for M independent samples each of size N. This leads to the identification of up to M near-optimal feasible solutions. Statistical confidence intervals, based on Shapiro (2003), are then derived on the quality of the near-optimal solutions found. An example of the application of this approach to a related network design problem is given in Santoso *et al*. (2005). The next sections present the approach proposed to generate a sample of environments Ω^N and the SAA solution procedure.

3.2 Environment sample generation

In order to formulate the SAA program, one must first generate a valid sample of environments Ω^N. Note that in order to select the sample $\left\{\omega^1,...,\omega^N\right\} \subset \Omega$, it is not necessary to use the probabilities $p(\omega), \omega \in \Omega$ explicitly. Since $p(\omega)$ must be derived from the probabilities $\theta_i,\ i \in I$, of signing contracts and VMI agreements, it is easier to construct sampled environments directly from these probabilities. Assuming that the customers decisions are taken independently of each other, to sample an environment $\omega \in \Omega^N$ we start by generating a pseudorandom set $\left\{u_i^c(\omega)\right\},\ c \in \cup_{o\in O}(PC_o \cup PV_o),\ i \in I_c$, of independent numbers uniformly distributed on the interval $\left[0;1\right]$. Using these numbers, the elements of the environment vector ω are then obtained with the following transformation:

$$\omega_i = \begin{cases} 1 & c(i) \in \cup_{o\in O}(PC_o \cup PV_o),\ i \in I_{c(i)}(\omega) \\ 0 & c(i) \in \cup_{o\in O}(PC_o \cup PV_o),\ i \notin I_{c(i)}(\omega) \\ 1 & \qquad c(i) \in \cup_{o\in O}(SC_o \cup SV_o) \end{cases}, \quad \text{where } I_c(\omega) = \left\{i \in I_c \left| \theta_i \geq u_i^c(\omega)\right.\right\}$$

The subsets $I_c(\omega) \subset I_c$ thus defined indicate which logistics policies would lead to signed contracts/agreements if implemented. Repeating the previous Monte Carlo sampling method *N* times yields the required sample of environments $\Omega^N = \left\{\omega^1,...,\omega^N\right\}$.

3.3 Sample average approximation

As explained at the end of section 3.1, the SAA approach involves the solution of program SAA(Ω^N) for *M* different samples of size *N*. This implies that *M* different near-optimal feasible designs could be obtained and the questions to answer are then: which design is the best and how close is it to the true optimum? To answer these questions, we need to obtain better estimates of the true value of the objective function of the solutions found through a Monte-Carlo evaluation based on a sample size N' much bigger than *N*. We also need to obtain statistical lower and upper bounds on the true value of the optimal solution of (SP). Let v^* and v^N be the optimal value of program (SP) and program SAA(Ω^N), respectively, and let $(\hat{\mathbf{X}}^N, \hat{\mathbf{Y}}^N, \hat{\mathbf{Z}}^N)$ be an optimal solution of SAA(Ω^N). Also, let us rewrite the SAA program as follows:

$$\max_{\mathbf{X}\geq\mathbf{0};\mathbf{Y},\mathbf{Z}\text{ bin}} \left\{ \hat{f}^N(\mathbf{X},\mathbf{Y},\mathbf{Z}) = \frac{1}{N}\sum_{\omega\in\Omega^N} Q(\mathbf{X},\mathbf{Y},\mathbf{Z},\omega) - \left(\mathbf{cX}+\mathbf{aY}+\mathbf{kZ}\right) \middle| \mathbf{A}\left[\mathbf{X},\mathbf{Z}\right] \leq \mathbf{b} \right\} \qquad \text{SAA}(\Omega^N)$$

where $Q(\mathbf{X},\mathbf{Y},\mathbf{Z},\omega)$ is the optimal value of the second stage linear program:

$$Q(\mathbf{X},\mathbf{Y},\mathbf{Z},\omega) = \max_{\mathbf{F}\geq\mathbf{0}} \left\{ \mathbf{qF} \middle| \mathbf{UF} = \mathbf{T}(\omega)\mathbf{Z},\ \mathbf{WF} = \mathbf{h} + \mathbf{V}\left[\mathbf{X},\mathbf{Y}\right] \right\}$$

It can be shown that the expected value of v^N is greater than or equal to v^* (Shapiro, 2003). This result is used to derive the required statistical upper bound. Also, since $(\hat{\mathbf{X}}^N, \hat{\mathbf{Y}}^N, \hat{\mathbf{Z}}^N)$ is a feasible solution of program (SP), we have $f(\hat{\mathbf{X}}^N, \hat{\mathbf{Y}}^N, \hat{\mathbf{Z}}^N) \leq v^*$. This is used to obtain the required statistical lower bound. From these observations, it is seen that a near-optimal design is found using the following procedure (a similar procedure is found in Santoso *et al.*, 2005):

<u>*Step 1*</u>: Generate *M* independent samples each of size *N*, $\left\{\omega_j^1,\ldots,\omega_j^N\right\} = \Omega_j^N$, $j=1,\ldots,M$ and solve SAA(Ω^N) for each sample *j*. Let v_j^N and $(\hat{\mathbf{X}}^N, \hat{\mathbf{Y}}^N, \hat{\mathbf{Z}}^N)$ be the corresponding optimal objective value and an optimal solution, respectively.

<u>*Step 2*</u>: Compute the statistical lower bound

$$\overline{v}_{N,M} = \frac{1}{M}\sum_{j=1}^{M} v_j^N$$

Since $\overline{v}_{N,M}$ is an unbiased estimator of $E(v_N)$, we have $\overline{v}_{N,M} \geq v^*$.

<u>*Step 3*</u>: Let *J* be the set of distinct solutions found with the samples $j=1,\ldots,M$.

For each distinct solution found, $(\hat{\mathbf{X}}_j^N, \hat{\mathbf{Y}}_j^N, \hat{\mathbf{Z}}_j^N), j \in J$, estimate its true objective function value $f(\hat{\mathbf{X}}_j^N, \hat{\mathbf{Y}}_j^N, \hat{\mathbf{Z}}_j^N)$ with the following approximation:

$$\tilde{f}^{N'}(\hat{\mathbf{X}}_j^N, \hat{\mathbf{Y}}_j^N, \hat{\mathbf{Z}}_j^N) = \frac{1}{N'} \sum_{\omega \in \Omega^{N'}} Q(\hat{\mathbf{X}}_j^N, \hat{\mathbf{Y}}_j^N, \hat{\mathbf{Z}}_j^N, \omega) - \left(\mathbf{c}\hat{\mathbf{X}}_j^N + \mathbf{a}\hat{\mathbf{Y}}_j^N + \mathbf{k}\hat{\mathbf{Z}}_j^N \right)$$

Note that the sample of size N' must be generated independently of the sample used to obtain $(\hat{\mathbf{X}}_j^N, \hat{\mathbf{Y}}_j^N, \hat{\mathbf{Z}}_j^N)$. For each $j \in J$, this step requires the solution of N' second-stage linear programs to obtain the optimal values $Q(\hat{\mathbf{X}}_j^N, \hat{\mathbf{Y}}_j^N, \hat{\mathbf{Z}}_j^N, \omega),\ \omega \in \Omega^{N'}$. Note that $\tilde{f}^{N'}(\hat{\mathbf{X}}_j^N, \hat{\mathbf{Y}}_j^N, \hat{\mathbf{Z}}_j^N)$ is an unbiased estimator of $f(\hat{\mathbf{X}}_j^N, \hat{\mathbf{Y}}_j^N, \hat{\mathbf{Z}}_j^N)$ (Shapiro, 2003) and thus, it is a statistical lower bound on v^*. The statistical bounds obtained can be used to compute an estimate of the optimality gap of solution $(\hat{\mathbf{X}}_j^N, \hat{\mathbf{Y}}_j^N, \hat{\mathbf{Z}}_j^N)$:

$$\mathbf{Gap}_{N,M,N'}(\hat{\mathbf{X}}_j^N, \hat{\mathbf{Y}}_j^N, \hat{\mathbf{Z}}_j^N) = \overline{v}_{N,M} - \tilde{f}^{N'}(\hat{\mathbf{X}}_j^N, \hat{\mathbf{Y}}_j^N, \hat{\mathbf{Z}}_j^N)$$

Step 4: Select the solution $(\hat{\mathbf{X}}_j^N, \hat{\mathbf{Y}}_j^N, \hat{\mathbf{Z}}_j^N), j \in J$, with the largest estimated objective function value $\tilde{f}^{N'}(\hat{\mathbf{X}}_j^N, \hat{\mathbf{Y}}_j^N, \hat{\mathbf{Z}}_j^N)$.

Having selected the best design, one can check its **Gap** to see if it is reasonable. If not, then additional samples must be used and/or the sample size *N* must be increased in order to get a better solution. Note that an expression for the variance of the above **Gap,** which can be used to compute a confidence upper bound for the **Gap**, is derived in Shapiro (2003).

4 Experimental Evaluation

4.1 Test case

In order to test the applicability and feasibility of the approach, we developed a realistic case, based on typical production-distribution network design problems encountered in the forest products industry. The characteristics of the case are summarized in Table 2. Most of the case data were taken from real lumber companies but the probabilities of getting the different contracts/agreements considered were randomly generated. The first stage design decisions specify the mission of the sawmills, the number and location of the warehouses and the logistics policies to implement. Since the case considers 21 potential logistics policies, the number of possible environments is about two millions, and model (SP) includes billions of second stage variables. Consequently, the need to use the Sample Average Approximation is clear, even for this moderate size case.

Product Families	19
Sawmills	3
Countries	2
Potential warehouses	7
Spot markets	4
Demand zones	16
Pre-signed contracts/agreements	2
Potential customers	13
Potential logistics policies	21
Possible environments	$2^{21} \approx 2\,000\,000$

Table 2: Test Case Characteristics

Contract Demand	**Price Difference** *($/unit)* (Contracts – Spot)	**Reference Price** *($/unit)*	
		High (440)	*Low (340)*
*40%*Capacity*	*40*	**# 1**	**# 3**
*20%*Capacity*	*40*	**# 2**	**# 4**
*40%*Capacity*	*55*	**#5** *(Average reference price (390))*	

Table 3: Test Case Instances

In order to test the solvability of the SAA model under extreme conditions, five instances of the case were created with different demand and price values, as described in Table 3. Our aim was also to understand the influence of demand and price differences, between the spot market and the contracts/agreements, on logistics policies and warehouse location decisions. Contracts/agreements become more interesting for the company when the price difference is high, which should lead to the implementation of more warehouses to support the selected logistics policies.

4.2 Computational results

Several sample sizes N and N' were tested in the experiments, in order to evaluate their impact on computational times and on the quality of the solutions obtained. The SAA models were solved with $M = 5$ independent samples, each of size $N = 5, 25, 50, 75, 100$. The number of variables and constraints in the models obtained for each sample size are given in Table 4. The mathematical programs were solved with CPLEX 9.0 on a 1.9 Ghz computer. The computational times observed are similar for the five case instances, but they increase exponentially with the sample size *N*. For example, Figure 4 presents the computational times (in seconds) obtained for the five samples generated for case instance #5.

N	**Variables**		**Constraints**	
	Binary	**Continuous**	**Equality**	**Inequality**
5	28	28 216	2 076	1 788
25	28	140 671	10 288	8 648
50	28	281 260	20 565	17 223
75	28	421 859	30 834	25 798
100	28	562 510	41 110	34 373

Table 4: SAA Model Statistics for Different Sample Sizes

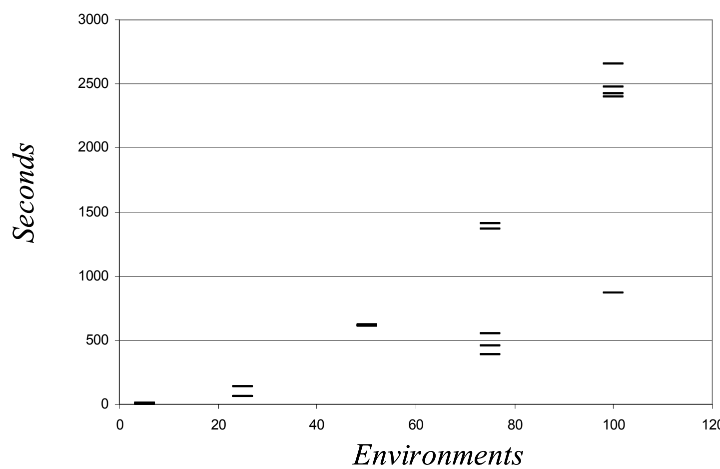

Figure 4: Computational Times (in seconds) for Instance #5

The analysis of results dissociates two clusters of case instances: those converging to a single optimal solution as *N* increases, and those for which a set of distinct solutions is obtained. The first cluster includes case instances #1, #3 and #4. For each of these cases, when $N \geq 25$, a single solution is obtained ($|J| = 1$), as illustrated in Figure 5 for case instance #1. The figure also shows that, as *N* increases, the value of the objective function for the 5 samples converges to the same value. Clearly, for these well behaved cases, no further analysis is required since a single solution is obtained. The second cluster composed of instances #2 and #5 is quite different. Indeed, several solutions are obtained, as illustrated in Figure 6 for case instance #5. The results of the application of the SAA procedure presented in section 4 to this case are provided in Table 5, for $M = 5$, $N = 75$, 100 and $N' = 200,\ 300,\ 400,\ 500$. For each distinct solution j, the objective function value approximation $\tilde{f}^{N'}(\hat{\mathbf{X}}_j^N, \hat{\mathbf{Y}}_j^N, \hat{\mathbf{Z}}_j^N)$ and the **Gap**, expressed in %, are reported. Moreover, for comparison purposes, the objective function value approximation $\tilde{f}^{N'}(\bar{\mathbf{X}}, \bar{\mathbf{Y}}, \bar{\mathbf{Z}})$ is also given for the solutions obtained, $(\bar{\mathbf{X}}, \bar{\mathbf{Y}}, \bar{\mathbf{Z}})$, when a deterministic version of the model is solved with the average demand (Average), and with the most probable environment demand (Probable).

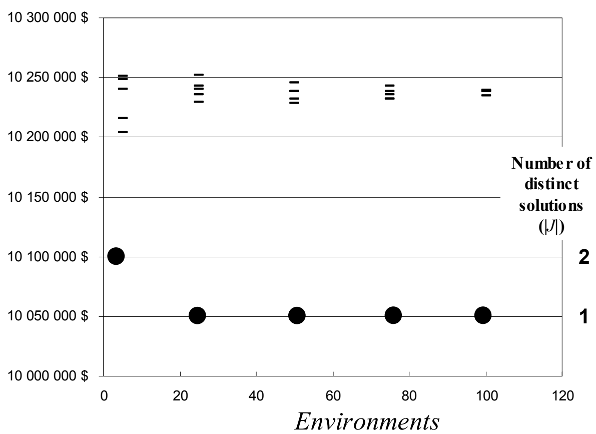

Figure 5: Results for Instance #1

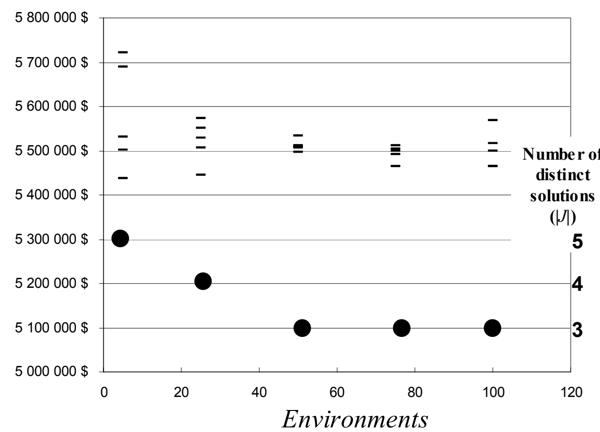

Figure 6: Results for Instance #5

The best design obtained for all the values of N and N' is the same (solution 3 for N=75 and solution 1 for N=100). In all these cases, the **Gap** is very small (not larger than 0,61%) which means that this solution is probably very good. This is comforting since it means that, at least for the cases considered, the Sample Average Approximation can be expected to give very good results even if relatively small sample sizes are used. Note also that, in this case, the solution obtained with the mean demand deterministic model is the same as the one obtained with the SAA model. This is not generally the case however and, in fact, there is no guarantee that the solution obtained with the average demand is a feasible solution of model (SP) because the expected demand is a fraction of the contracts/agreements demand. Lastly, note that the solution obtained when a deterministic model with the most probable environment is used is not very good. This suggest that using the SAA method gives solutions which can be much better than those obtained with the type of deterministic models found in the literature.

(M=5)		SAA Method ($)						Deterministic	
		$N=75$			$N=100$			Probable	Average
Sample (j)		1	2	3	1	2	3	1	1
Duplicates			(100,3)	(100,1), (A1)	(75,3), (A1)		(75,2)		(75,3), (100,1)
$N'=200$	$\tilde{f}^N$	5 466 583	5 486 456	5 494 382	5 494 382	5 474 509	5 486 456	5 073 303	5 494 382
	Gap	0,50%	0,13%	0,01%	0,15%	0,51%	0,29%		
$N'=300$	$\tilde{f}^N$	5 468 001	5 493 454	5 496 373	5 496 373	5 470 920	5 493 454	5 067 651	5 496 373
	Gap	0,47%	0,00%	0,05%	0,11%	0,58%	0,16%		
$N'=400$	$\tilde{f}^N$	5 461 504	5 480 940	5 494 136	5 494 136	5 474 700	5 480 940	5 085 692	5 494 136
	Gap	0,59%	0,23%	0,01%	0,15%	0,51%	0,39%		
$N'=500$	$\tilde{f}^N$	5 461 779	5 477 564	5 484 727	5 484 727	5 468 941	5 477 564	5 071 398	5 484 727
	Gap	0,58%	0,29%	0,16%	0,32%	0,61%	0,45%		

Table 5: SAA Procedure Results for Instance #5

Instances	Average number of logistics policies selected (21 potential)	Average number of warehouse selected (7 potential)
# 1	4	2
# 2	5.5	2
# 3	4	2
# 4	4	2
# 5	11	3.8

Table 6: Designs Obtained for *N*=100 and *M*=5

The designs obtained for the 5 case instances studied are summarized in Table 6, for samples size of $N=100$. A close observation of these results confirms that our initial intuition was correct. The warehouses and policies selected for the four first case instances are roughly the same. This means that the solution is not very sensitive to changes in demand and in reference prices. However, the solution obtained for instance #5 involves the selection of a much higher number of policies and warehouses, which implies that the optimal design is very sensitive to the difference in the price of products between the spot market and contracts/agreements.

4.3 Virtu@l Lumber case

The design approach developed in this paper was also applied to a case involving a much more elaborated production-distribution network. It emulates a typical lumber industry company in the province of Quebec in Canada, and the case, model and experimental result details are provided in Vila et al. (2006), which also show how the approach can be used for strategic decision making. Extensive experiments were made to study the impact of price differentials between market segments on the production-distribution network structure and on the number of contract signatures. Different forest management policies and business acquisition and rationalization options were also studied. The problems were solved with the SAA method using five independent samples of twenty-five environments ($M = 5$, $N = 25$). In order to determine the best design, the true value of the objective function was estimated with samples of size $N' = 100$.

The models solved in these runs include about 900 000 continuous variables, 200 binary variables and 250 000 constraints, and the computational times required to solve them with CPLEX 9.0 ranged between 321 and 1 670 minutes. The optimization process was stopped when the computation time reached 100 000 seconds, which means that all the SAA problems were not necessarily solved to optimality. The SAA gap of the problems solved ranged between 0.1 % and 4.5 %. In general, the design obtained for the production sub-network was the same for the five (M = 5) samples used. However, for the distribution sub-network, 1 to 3 different designs were obtained and, in most cases, 5 different logistics policy vectors were obtained. This indicates that the design variables 'closest' to the market are more sensitive to the scenarios generated. In all cases, the solutions suggested by the model made good business sense. Note also that for these more realistic cases, the designs obtained when using an average demand model were never the best and they were often unfeasible for some of the scenarios generated, which clearly shows the advantage of the stochastic programming approach proposed.

5 Conclusion

The production-distribution network design methodology proposed in this paper takes market considerations into account to obtain designs improving the competitive position of the company or companies involved. Furthermore, the two-stage stochastic programming model proposed and the Monte-Carlo sampling method used to solve the model lead to robust designs which can be expected to perform well under most possible future business environments. The experiments made with the model show that good results can be obtained even if a relatively small environment sample size is used, at least for the lumber industry context on which our test cases were based. For these cases, the model proposed is easy to solve for moderate size problems but much more difficult to solve to optimality for large cases. These first applications of the model to realistic cases from the lumber industry are very promising, but computational testing in other industrial contexts would be required to demonstrate the general value of the model. Future studies will involve the development of tailor-made acceleration techniques and heuristics to solve very large network design models in a reasonable time.

References

[1] Arntzen, B., G. Brown, T. Harrison and L. Trafton, 1995, Global Supply Chain Management at Digital Equipment Corporation, *Interfaces* 21, 1, 69-93.

[2] Ben-Akiva M. and S. Lerman, 1985, *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press.

[3] Birge, J. and F. Louveaux, 1997, *Introduction to Stochastic Programming*, Springer.

[4] Cohen, M., M. Fisher and R. Jaikumar, 1989, International Manufacturing and Distribution Networks: A Normative Model Framework, in K. Ferdows (ed), *Managing International Manufacturing*, Elsevier, 67-93.

[5] Cohen, M. and H. Lee, 1989, Resource Deployment Analysis of Global Manufacturing and Distribution Networks, *Journal of Manufacturing and Operations Management* 2, 81-104.

[6] Cohen, M. and S. Moon, 1990, Impact of Production Scale Economies, Manufacturing Complexity, and Transportation Costs on Supply Chain Facility Networks, *Journal of Manufacturing and Operations Management* 3, 269-292.

[7] Cohen, M. and S. Moon, 1991, An integrated plant loading model with economies of scale and scope, *European Journal of Operational Research* 50, 266-279.

[8] Dogan K. and M. Goetschalckx, 1999, A Primal Decomposition Method for the Integrated Design of Multi-Period Production-Distribution Systems, *IIE Transactions* 31, 1027-1036.

[9] Eppen, G., R. Kipp Martin and L. Schrage, 1989, A Scenario Approach to Capacity Planning, *Operations Research* 37, 517-527.

[10] Fandel, G. and M. Stammen, 2004, A General Model for Extended Strategic Supply Chain Management with Emphasis on Product Life Cycles Including Development and Recycling, *International Journal of Production Economics* 89, 293-308.

[11] Hill, T., 1994, *Manufacturing Strategy*, Irwin.

[12] Geoffrion, A. and G. Graves, 1974, Multicommodity Distribution System Design by Benders Decomposition, *Management Science* 20, 822-844.

[13] Geoffrion, A. and R. Powers, 1995, 20 Years of Strategic Distribution System Design: Evolutionary Perspective, *Interfaces* 25, 5, 105-127.

[14] Louviere, J., D. Hensher and J. Swait, 2000, *Stated Choice Methods*, Cambridge University Press.

[15] Martel, A. and U. Vankatadri, 1999, Optimizing Supply Network Structures Under Economies of Scale, *IEPM Conference Proceedings*, Glasgow, Book 1, 56-65.

[16] Martel, A., 2005, The design of production-distribution networks: A mathematical programming approach, in J. Geunes and P. Pardalos (eds.), *Supply Chain Optimization*, Springer, 265-306.

[17] Mazzola, J. and R. Schantz, 1997, Multiple-Facility Loading Under Capacity-Based Economies of Scope, *Naval Research Logistics* 44, 229-256.

[18] Mazzola, J. and A.W. Neebe, 1999, Lagrangian-Relaxation-Based Solution Procedures for a Multiproduct Capacitated Facility Location Problem with Choice of Facility Type, *European Journal of Operational Research* 115, 285-299.

[19] Pirkul, H. and V. Jayaraman, 1996, Production, Transportation, and Distribution Planning in a Multi-Commodity Tri-Echelon System, *Transportation Science* 30, 291-302.

[20] Paquet, M., A. Martel and G. Desaulniers, 2004, Including Technology Selection Decisions in Manufacturing Network Design Models, *International Journal of Computer Integrated Manufacturing* 17, 117-125.

[21] Rosenfield, D., R. Shapiro and R. Bohn, 1985, The implications of cost-service trade-offs on industry logistics structures, *Interfaces* 15, 47-59.

[22] Santoso, T., S. Ahmed, M. Goetschalckx and A. Shapiro, 2005, A stochastic programming approach for supply chain network design under uncertainty, *European Journal of Operational Research* 167, 96-115.

[23] Shapiro, A., 2003, Monte Carlo Sampling Methods, in A. Ruszczynski and A. Shapiro (eds), *Stochastic Programming*, Handbooks in Operations Research and Management Science, 10, Elsevier.

[24] Shapiro, J., V. Singhal and S. Wagner, 1993, Optimizing the Value Chain, *Interfaces* 23, 2, 102-117.

[25] Shapiro, J., 2001, *Modeling the Supply Chain*, Brooks/Cole Publishing Company.

[26] Tsay A., S. Nahmias and N. Agrawal,1999, Modeling Supply Chain Contracts : A Review, *Quantitative Models for Supply Chain Management*, Kluwer Academic Publishers.

[27] Verter, V. and C. Dincer, 1992, An integrated evaluation of facility location, capacity acquisition, and technology selection for designing global manufacturing strategies, *European Journal of Operational Research* 60, 1-18.

[28] Verter, V. and C. Dincer, 1995, Facility Location and Capacity Acquisition: An Integrated Approach, *Naval Research Logistics* 42, 1141-1160.

[29] Vidal, C. and M. Goetschalckx, 1997, Strategic Production-Distribution Models: A Critical Review with Emphasis on Global Supply Chain Models, *European Journal of Operational Research* 98, 1-18.

[30] Vidal, C. and M. Goetschalckx, 2001, A Global Supply Chain Model with Transfer Pricing and Transportation Cost Allocation, *European Journal of Operational Research* 129, 134-158

[31] Vila, D., R. Beauregard and A. Martel, The Strategic Design of Forest Industry Supply Chains, Working Paper DT-2005-RB-2, Centor, U. Laval, 2006.

Appendix

Proposition: Any optimal solution of program SAA(Ω^N) is such that, for step $\lambda \in \Lambda_{pm}$ of the spot market price function:

$$\sum_{p' \in SP^p} \sum_{s \in S_m^i} F_{pp'sd\lambda}(\omega) > 0 \Rightarrow \sum_{p' \in SP^p} \sum_{s \in S_m^i} F_{pp'sd\lambda'}(\omega) = \kappa_d (X_{pm\lambda'} - X_{pm(\lambda'-1)}), \quad \lambda' < \lambda$$

Proof: Let's consider an **optimal** solution $(\mathbf{X}^\lozenge, \mathbf{Y}^\lozenge, \mathbf{Z}^\lozenge, \mathbf{F}^\lozenge(\omega))$ of SAA (Ω^N). Assume that $\exists \omega^\circ \in \Omega^N,\ o^\circ \in O,\ m^\circ \in SM_o,\ p^\circ \in P_m,\ d^\circ \in D_{pm},\ \lambda^\circ \in \Lambda_{pm}$ such that

$\sum_{p' \in SP^{p^\circ}} \sum_{s \in S_m^i} F^\lozenge_{p^\circ p'sd^\circ\lambda^\circ}(\omega^\circ) > 0$ and $\sum_{p' \in SP^p} \sum_{s \in S_m^i} F^\lozenge_{p^\circ p'sd^\circ\lambda'^\circ}(\omega^\circ) < \kappa_{d^\circ}(X_{p^\circ m^\circ \lambda'^\circ} - X_{p^\circ m^\circ(\lambda'^\circ-1)}),\ \lambda'^\circ < \lambda^\circ$

Then, clearly, $\exists p'^\circ \in SP^{p^\circ},\ s^\circ \in S^i_{m^\circ}$ with $F^\lozenge_{p^\circ p'^\circ s^\circ d^\circ \lambda^\circ}(\omega^\circ) > 0$. Now, let's construct the new feasible solution $(\mathbf{X}^\ell, \mathbf{Y}^\ell, \mathbf{Z}^\ell, \mathbf{F}^\ell(\omega)) = (\mathbf{X}^\lozenge, \mathbf{Y}^\lozenge, \mathbf{Z}^\lozenge, \left[F^\lozenge_{pss'}(\omega), F^\ell_{pp'sd\lambda}(\omega), F^\lozenge_{psi}(\omega)\right])$ with

$$F^\ell_{pp'sd\lambda}(\omega) = F^\lozenge_{pp'sd\lambda}(\omega),\ \lambda \neq \lambda^\circ, \lambda'^\circ$$

$$F^\ell_{p^\circ p'^\circ s^\circ d^\circ \lambda'^\circ}(\omega^\circ) = F^\lozenge_{p^\circ p'^\circ s^\circ d^\circ \lambda'^\circ}(\omega^\circ) + \max(\kappa_{d^\circ}(X_{p^\circ m^\circ \lambda'^\circ} - X_{p^\circ m^\circ(\lambda'^\circ-1)}) - \sum_{p' \in SP^{p^\circ}} \sum_{s \in S^i_{m^\circ}} F^\lozenge_{p^\circ p'sd^\circ\lambda'^\circ}(\omega^\circ); F^\lozenge_{p^\circ p'^\circ s^\circ d^\circ \lambda^\circ}(\omega^\circ))$$

$$F^\ell_{p^\circ p'^\circ s^\circ d^\circ \lambda^\circ}(\omega^\circ) = F^\lozenge_{p^\circ p'^\circ s^\circ d^\circ \lambda^\circ}(\omega^\circ) - \max(\kappa_{d^\circ}(X_{p^\circ m^\circ \lambda'^\circ} - X_{p^\circ m^\circ(\lambda'^\circ-1)}) - \sum_{p' \in SP^{p^\circ}} \sum_{s \in S^i_{m^\circ}} F^\lozenge_{p^\circ p'sd^\circ\lambda'^\circ}(\omega^\circ); F^\lozenge_{p^\circ p'^\circ s^\circ d^\circ \lambda^\circ}(\omega^\circ))$$

by reallocating the step λ° flow values on the price step functions.

If $(\mathbf{X}^\lozenge, \mathbf{Y}^\lozenge, \mathbf{Z}^\lozenge, \mathbf{F}^\lozenge(\omega))$ is optimal, the difference

$$\frac{1}{N}\sum_{\omega \in \Omega_N} \mathbf{qF}^\lozenge(\omega) - \left(\mathbf{cX}^\lozenge + \mathbf{aY}^\lozenge + \mathbf{kZ}^\lozenge\right) - \left(\frac{1}{N}\sum_{\omega \in \Omega_N} \mathbf{qF}^\ell(\omega) - \left(\mathbf{cX}^\ell + \mathbf{aY}^\ell + \mathbf{kZ}^\ell\right)\right)$$

$$= \frac{1}{N} P_{p^\circ m^\circ \lambda'^\circ}(F^\lozenge_{p^\circ p'^\circ s^\circ d^\circ \lambda'^\circ}(\omega^\circ) - F^\ell_{p^\circ p'^\circ s^\circ d^\circ \lambda'^\circ}(\omega^\circ)) + \frac{1}{N} P_{p^\circ m^\circ \lambda^\circ}((F^\lozenge_{p^\circ p'^\circ s^\circ d^\circ \lambda^\circ}(\omega^\circ) - F^\ell_{p^\circ p'^\circ s^\circ d^\circ \lambda^\circ}(\omega^\circ))$$

$$= \frac{1}{N}(P_{p^\circ m^\circ \lambda^\circ} - P_{p^\circ m^\circ \lambda'^\circ}) \max(\kappa_{d^\circ}(X_{p^\circ m^\circ \lambda'^\circ} - X_{p^\circ m^\circ(\lambda'^\circ-1)}) - \sum_{p' \in SP^{p^\circ}} \sum_{s \in S^i_{m^\circ}} F^\lozenge_{p^\circ p'sd^\circ\lambda'^\circ}(\omega^\circ); F^\lozenge_{p^\circ p'^\circ s^\circ d^\circ \lambda^\circ}(\omega^\circ))$$

should be positive. However, by *assumption*, we have

$$\max(\kappa_{d^\circ}(X_{p^\circ m^\circ \lambda'^\circ} - X_{p^\circ m^\circ(\lambda'^\circ-1)}) - \sum_{p' \in SP^{p^\circ}} \sum_{s \in S^i_{m^\circ}} F^\lozenge_{p^\circ p'sd^\circ\lambda'^\circ}(\omega^\circ); F^\lozenge_{p^\circ p'^\circ s^\circ d^\circ \lambda^\circ}(\omega^\circ)) > 0$$

and by *construction* of the price function we have $0 > (P_{p^\circ m^\circ \lambda^\circ} - P_{p^\circ m^\circ \lambda'^\circ})$, which implies that:

$$\frac{1}{N}\sum_{\omega \in \Omega_N} \mathbf{qF}^\lozenge(\omega) - \left(\mathbf{cX}^\lozenge + \mathbf{aY}^\lozenge + \mathbf{kZ}^\lozenge\right) - \left(\frac{1}{N}\sum_{\omega \in \Omega_N} \mathbf{qF}^\ell(\omega) - \left(\mathbf{cX}^\ell + \mathbf{aY}^\ell + \mathbf{kZ}^\ell\right)\right) < 0$$

Hence, $(\mathbf{X}^\lozenge, \mathbf{Y}^\lozenge, \mathbf{Z}^\lozenge, \mathbf{F}^\lozenge(\omega))$ is not optimal and the proposition is true. □