Math. Program., Ser. B 108, 355–394 (2006)

Digital Object Identifier (DOI) 10.1007/s10107-006-0715-7

Vikas Goel · Ignacio E. Grossmann

A Class of stochastic programs with decision dependent uncertainty ⋆

Received: August 25, 2004 / Accepted: April 2, 2005
Published online: April 25, 2006 – © Springer-Verlag 2006

Abstract. We address a class of problems where decisions have to be optimized over a time horizon given that the future is uncertain and that the optimization decisions influence the time of information discovery for a subset of the uncertain parameters. The standard approach to formulate stochastic programs is based on the assumption that the stochastic process is independent of the optimization decisions, which is not true for the class of problems under consideration. We present a hybrid mixed-integer disjunctive programming formulation for the stochastic program corresponding to this class of problems and hence extend the stochastic programming framework. A set of theoretical properties that lead to reduction in the size of the model is identified. A Lagrangean duality based branch and bound algorithm is also presented.

Nomenclature

Notation used in $(CAPEXP)$

Indices

i	Unit
j	Stream
t, τ	Time period
s, s'	Scenario

Optimization Variables

$b_{i,t}^s$	Whether or not unit i is operated in time period t, scenario s
$y_{i,t}^{exp,s}$	Whether or not unit i is expanded in time period t, scenario s
$y_{i,t}^{QE,s}$	Expansion in capacity of unit i in time period t, scenario s
$y_{j,t}^{rate,s}$	Flow rate of stream j in time period t, scenario s (decision variable; see Figure 2)
$x_t^{sales,s}$	Sales of chemical A in time period t, scenario s
$x_t^{purch,s}$	Purchases of chemical A in time period t, scenario s
$w_{j,t}^{rate,s}$	Flow rate of stream j in time period t, scenario s (state variable; see Figure 2)
$w_{i,t}^{Q,s}$	Capacity of unit i in time period t, scenario s

V. Goel, I. E. Grossmann: Department of Chemical Engineering, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213, USA.
e-mail: vikas.goel@exxonmobil.com, grossmann@cmu.edu

⋆ Financial support from the ExxonMobil Upstream Research Company is gratefully acknowledged.

$w_t^{inv,s}$ Inventory of chemical A at the end of time period t, scenario s
$w^{cost,s}$ Total costs in scenario s
$Z_t^{s,s'}$ Whether or not scenarios s, s' are indistinguishable after operation in time period t

Parameters

p^s Probability of scenario s
α_t Purchasing price for chemical A in time period t
β_t Sales price for chemical A in time period t
γ_t Cost of maintaining inventory of chemical A in time period t
δ_t Duration of time period t
θ_i^s Yield of unit i in scenario s
ξ_t^s Demand of chemical A in time period t, scenario s
$FO_{i,t}$ Fixed operating cost for unit i in time period t
$VO_{j,t}^y$ Variable operating cost corresponding to variable $y_{j,t}^{rate,s}$
$VO_{j,t}^w$ Variable operating cost corresponding to variable $w_{j,t}^{rate,s}$
$FE_{i,t}$ Fixed expansion cost for unit i in time period t
$VE_{i,t}$ Variable expansion cost for unit i in time period t
$U_{(\cdot)}^{(\cdot)}$ Upper bounds
$L_{(\cdot)}^{(\cdot)}$ Lower bounds

Notation used in $(SIZES)$

Indices

i Size
t, τ Time period
s, s' Scenario

Optimization Variables

$b_{i,t}^s$ Whether or not size i is produced in time period t, scenario s
$y_{i,t}^s$ Number of units of size i produced in time period t, scenario s
$x_{i,i',t}^s$ Number of units of size i used to satisfy demand of size i' in time period t, scenario s
$w_{i,t}^s$ Inventory of size i at the end of time period t, scenario s
$Z_t^{s,s'}$ Whether or not scenarios s, s' are indistinguishable after production in time period t

Parameters

p^s Probability of scenario s
θ_i^s Variable production cost for size i in scenario s
$\xi_{i,t}^s$ Demand of size i in time period t, scenario s
α Total production capacity
β Total inventory capacity
σ Fixed cost for producing a specific size in given time period
ρ Unit substitution cost
μ Unit inventory cost per time period

$U_{(\cdot)}^{(\cdot)}$ Upper bounds
$L_{(\cdot)}^{(\cdot)}$ Lower bounds

1. Introduction

Stochastic programming deals with the problem of making optimal decisions in the presence of uncertainty. In stochastic programs, the uncertainty is represented by probability distributions and the interaction between the stochastic and decisions processes is modeled so that the decision-maker has the option of adjusting the decisions based on how the uncertainty unfolds. From the modeling perspective, most previous work in the stochastic programming literature deals with problems with *exogenous uncertainty* (Jonsbraten [11]) where the optimization decisions cannot influence the stochastic process.

Pflug [16] was the first to address the case with *endogenous uncertainty* where the underlying stochastic process depends on the optimization decisions. Previous work on this class of uncertainty is limited to a few papers only. Since this paper deals with endogenous uncertainty, we only review the previous work in the stochastic programming literature on this type of uncertainty. To motivate the need for this paper, we also present brief descriptions of some real world problems with this type of uncertainty. Reviews of previous work on problems with exogenous uncertainty can be found in Birge [3], Sahinidis [19] and Schultz [20].

The project decisions can influence the stochastic process in at least two ways. On one hand, the decision-maker may cause alteration of the probability distribution by making one possibility more likely than the other. On the other hand, the decision-maker may not directly affect the probability distributions but could act to get more accurate information by resolving the uncertainty (partially). The difference is that while in the first case the decision-maker can force one possibility to become more probable, in the second case the decision-maker can only become more sure as to which possibility may occur in future.

Viswanath et al. [22] address an instance of the first type of endogenous uncertainty where optimization decisions can influence the probability distribution. They consider a two-stage network traversal problem where each arc is associated with a probability that represents the probability of the arc being available for traversal after some disaster. In the first stage, investments are made to increase the probabilities associated with some of the arcs. This is followed by a random event (a disaster) which renders some of the arcs unavailable for traversal. In the second stage, a path from the source to the destination has to be traversed using the available arcs. The aim is to invest in the arcs so that the expected shortest path length from the source to the destination is minimized. This problem arises in planning disaster relief between cities with the possibility that some of the inter-connecting routes may become unusable due to a disaster.

Ahmed [1] presents more examples relating to network design, server selection and facility location where the decision-maker can influence the probability distributions. The author presents a 0-1 hyperbolic programming formulation and an exact solution algorithm for single stage problems with discrete decisions.

The gas field development planning problem (Goel and Grossmann [6]) is a real world example of the second type of endogenous uncertainty where the optimization decisions control the resolution of uncertainty. In this problem, a set of fields (reservoirs of gas) are available for production. The size and quality of the reserves of these fields are uncertain and the uncertainty in a field is resolved only when a facility is installed at that field. Thus, the investment decisions control the times when the uncertainties will be resolved. Therefore, apart from considering the large capital expenditures (over US $100 Million) and revenues associated with investment at a field, it is also important to consider the potential of obtaining valuable information as a result of the investment. This information could lead to "better" decisions in the future.

A similar problem is the capacity expansion of process networks under yield uncertainty. In this problem, an existing network of processing units can be expanded by installing units that are based on new technology. The yields (or productivities) of these units are uncertain and the uncertainty in a unit is resolved only after that unit is installed and operated in the existing conditions. Thus, the investment decisions determine the times when the uncertainties will be resolved. We use this problem in section 8 to illustrate that when the value of information is sufficiently high, it may be optimal for the decision-maker to first resolve the uncertainty by making small investments and then make the major investments based on the observations.

Another instance of this type of endogenous uncertainty arises in the multistage network interdiction problem. In each stage, the interdictor interdicts some of the nodes followed by which the operator tries to traverse the network along the shortest available path. The exact network structure is unknown to the interdictor, but various possibilities are postulated through a set of scenarios. In each stage, the uncertainty is (partially) resolved based on the path taken by the operator, which is implicitly determined by the interdiction decisions. Thus, the aim of the interdictor is to interdict the nodes such that the most "valuable" information is obtained and the objective maximized.

Jonsbraten et al. [12] first addressed problems with endogenous uncertainty where project decisions give more accurate information by resolving the uncertainty. The authors present an implicit enumeration algorithm for this class of problems. Results for two-stage problems are also presented. Held and Woodruff [9] present heuristic solution methods for the multistage network interdiction problem. Both these papers assume that every resolution of uncertainty excludes at least one realization or scenario from the set of future possibilities. Jonsbraten [11] addresses a variant of the oil (or gas) field problem where investment decisions lead to resolution of uncertainty but may not exclude any of the scenarios from the set of future possibilities. The author proposes an implicit enumeration algorithm where the resolution of uncertainty is modeled using a Bayesian approach.

Goel and Grossmann [6] used the gas field problem to illustrate an approach for formulating rigorous stochastic programs for problems with the second type of endogenous uncertainty. In this approach, the interaction between the decisions and the resolution of uncertainty is captured through a disjunctive formulation of the non-anticipativity constraints. The authors also present a heuristic algorithm to solve the gas field problem.

In this paper, we generalize the above approach to problems that have both exogenous and endogenous uncertainties. We consider the second type of endogenous uncertainty where the project decisions determine the times when the uncertainties in various parameters will be resolved. This paper is organized as follows. In section 2 we present a brief background on stochastic programming formulations with exogenous uncertainty. Next we present a generic description of the broad class of problems under consideration. Sections 4 and 5 explain the notation and the proposed stochastic program, respectively. Section 6 presents theoretical properties that lead to reduction in the dimensionality of the model. In section 7, we present a Lagrangean duality based branch and bound algorithm to solve the proposed model. Finally, results for two applications belonging to the class of problems under consideration are presented in section 8.

2. Background

We restrict the scope of this paper to problems where the uncertainty can be represented by discrete probability distributions and the time horizon is represented by a discrete set of time periods. For such problems, the stochastic process can be represented by a *scenario tree*. Each node represents a possible information state while each arc emanating from a node for time period t represents a possible transition to a node for time period $t+1$. Each arc is associated with a transition probability and multiple arcs emanating from a node for time period t represent multiple possibilities for transition and hence, that uncertainty in some parameter(s) will be resolved at the end of time period t. In a scenario tree, a path from the root node to a leaf node represents a *scenario*. A scenario represents one possible combination of values for all uncertain parameters. The probability of a scenario is the probability of reaching the corresponding leaf node from the root node.

Fig. 1(a) represents the scenario tree for a problem with two uncertain parameters ξ_1, ξ_2, and three time periods. Possible realizations for both parameters include H ("High") and L ("Low") where both realizations are equally probable. The uncertainties in ξ_1 and ξ_2 are resolved at the end of the first and second time periods, respectively. The scenario tree has four scenarios, each with probability equal to 0.25.

Ruszczynksi [18] illustrates an alternative representation of scenario trees where each scenario is represented by a set of unique nodes (Fig. 1(b)). If the nodes for scenarios s, s' in time period t correspond to the same information state (represented by horizontal dotted lines linking the nodes in Fig. 1(b)), the two scenarios are said to be *indistinguishable* in time period t. Scenarios s, s' are indistinguishable at a given instant of time if they are identical in realizations for all uncertain parameters in which uncertainty has been resolved up till that time. The concept of indistinguishability is central to the non-anticipativity (see Rockafellar and Wets [17], Ruszczynksi [18]) based approach to formulating a stochastic program.

(SSP) is a "standard" stochastic program (Jonsbraten et al. [12]) for a linear problem with T time periods and scenario tree $\mathcal{S}$.

$$\text{(SSP)}\ \min \sum_s p^s \sum_t c_t^s x_t^s \tag{1a}$$

$$\text{s.t.} \quad \sum_{\tau \le t} A_{\tau,t}^s x_\tau^s \le a_t^s \quad \forall (t, s) \tag{1b}$$

$$x_t^s \in \mathcal{X}_t^s \quad \forall (t, s) \tag{1c}$$

$$x_t^s = x_t^{s'} \quad \forall (s, s', t) \in \mathcal{N}_{\mathcal{S}}^e \tag{1d}$$

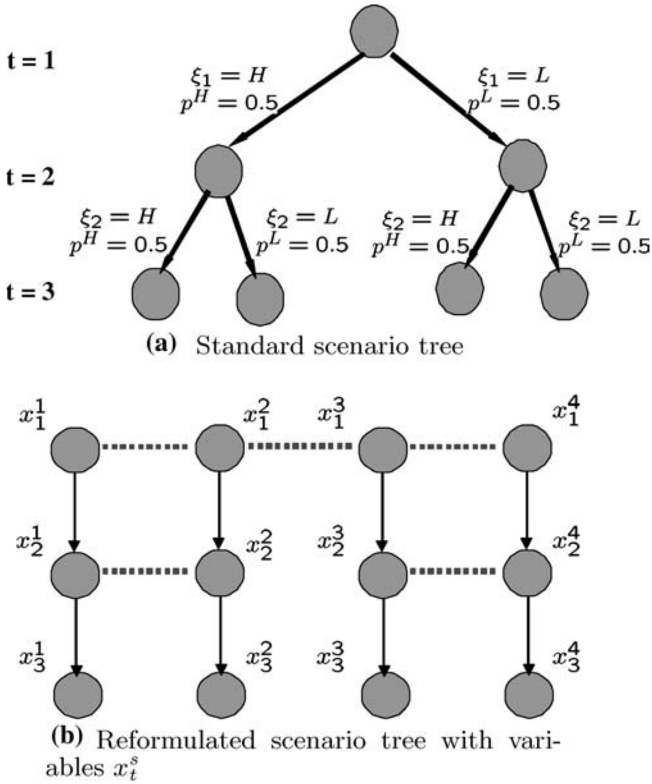


Fig. 1. Equivalent scenario trees

Parameter p^s represents the probability of scenario s while variables x_t^s represent decision variables for time period t in scenario s. (1a) represents the objective of minimizing the expectation of some economic criterion. Constraint (1b) represents single-period and period-linking constraints for a particular scenario which are characteristic of any multi-period model. Constraint (1c) represents integrality and bound restrictions on variables x_t^s. $\mathcal{N}_{\mathcal{S}}^e$ represents the set of tuples (s, s', t) such that scenarios s and s' are indistinguishable in time period t for scenario tree $\mathcal{S}$. The *non-anticipativity* or *implementability* constraints (1d) link decisions for different scenarios. These constraints state that if scenarios s, s' are indistinguishable in time period t then decisions for s, s' in t should be the same. In other words, decisions cannot be based on knowledge that will be revealed in the future.

When the uncertainty is of exogenous nature, the probabilities p^s and the set $\mathcal{N}_{\mathcal{S}}^e$ are independent of the optimization variables. Thus, these are inputs to the optimization model. However, probabilities p^s have to be treated as optimization variables if the optimization decisions influence the probability distribution. On the other hand, the scenario tree and hence the set $\mathcal{N}_{\mathcal{S}}^e$ depends on the decisions (see Jonsbraten et al. [12], Goel and Grossmann [6]) if the optimization decisions influence the time when the uncertainty is resolved. In this case, the set $\mathcal{N}_{\mathcal{S}}^e$ is not known a priori. We generalize the approach of Goel and Grossmann [6] to problems with both exogenous and endogenous uncertainty by formulating the inter-dependence of $\mathcal{N}_{\mathcal{S}}^{e}$ and the optimization variables as a disjunctive program.

3. Generic problem description

In the class of problems under consideration, the time horizon is represented by the discrete set of time periods $\mathcal{T} = \{1, 2, \ldots, T\}$. ξ_t represents the vector of exogenous uncertain parameters associated with time period $t \in \mathcal{T}$. The uncertainty in ξ_t will be resolved automatically in time period t. Ξ represents the discrete set of possible realizations for vector $\xi = (\xi_1, \xi_2, \ldots, \xi_T)$.

Set $\mathcal{I} = \{1, 2, \ldots, I\}$ represents the set of "sources" of endogenous uncertainty while θ_i represents the endogenous uncertain parameter associated with source $i \in \mathcal{I}$. The discrete set of possible realizations for θ_i is represented by Θ_i. The resolution of uncertainty in θ_i depends on binary decision variables $b_{i,t}$. Specifically, the uncertainty in θ_i will be resolved in time period t if binary decision $b_{i,t} = 1$ and $b_{i,\tau} = 0 \quad \forall \tau < t$. Note that parameters θ_i represent intrinsic properties of source i and are independent of time period. Besides decisions represented by variables $b_{i,t}$, other decisions to be made in time period t are represented by variables y_t and x_t together.

The sequence of events in each time period is as follows. Decisions y_t and $b_{i,t}$ are implemented at the beginning of time period t. This is followed by the resolution of uncertainty in the exogenous parameters ξ_t and in the endogenous parameter θ_i for source i if $b_{i,t} = 1$ and $b_{i,\tau} = 0 \quad \forall \tau < t$. Finally, decisions x_t are implemented at the end of the time period.

In general, variables $b_{i,t}$ may represent investment or operation decisions associated with source i. In the gas field problem considered by Goel and Grossmann [6], these variables represent whether or not investment is made at field i in time period t. The uncertainty associated with a field is resolved in time period t only if investment is carried out at that field in time period t while no investments have been made at that field in the past.

Note that for ease of exposition, we assume that there is only one endogenous uncertain parameter associated with source i for all $i \in \mathcal{I}$. Thus, θ_i is a *scalar* for all $i \in \mathcal{I}$. At the end of section 6 we describe how our approach extends to the more general case where θ_i may be a vector for some $i \in \mathcal{I}$.

4. Notations and definitions

In order to make the following discussion more comprehensible, we first explain the notation and definitions used in this paper. Each scenario in this problem corresponds to one possible realization for the vector $(\xi_1, \xi_2, \ldots, \xi_T, \theta_1, \theta_2, \ldots, \theta_I)$. We assume that the set of scenarios corresponds to $\Xi \times (\times_{i \in \mathcal{I}} \Theta_i)$, *i.e.*, for any realization of the vector of exogenous parameters, $\xi = (\xi_1, \xi_2, \ldots, \xi_T)$, the set of scenarios includes scenarios corresponding to all possible combinations of realizations for the endogenous parameters. Individual scenarios are indexed as $s \in \mathcal{S}$, where $\mathcal{S} = \{1, 2, \ldots, S\}$ represents the set of indices corresponding to all the scenarios. Note that we will use index s to refer to the corresponding scenario. Further, θ_i^s and ξ_t^s will represent the realizations of θ_i and ξ_t, respectively, in scenario s.

For scenarios $s, s' \in \mathcal{S}$, the set $\mathcal{D}\left(s, s'\right) = \left\{i | i \in \mathcal{I}, \theta_i^s \neq \theta_i^{s'}\right\}$ represents the set of sources of endogenous uncertainty that distinguish scenarios s and s'. The symbol $|\mathcal{D}\left(s, s'\right)|$ represents the cardinality of this set. In general, the set $\mathcal{D}(\cdot, \cdot)$ satisfies $0 \leq |\mathcal{D}\left(s, s'\right)| \leq I$ for all $s, s' \in \mathcal{S}$, where I is the number of sources of endogenous uncertainty. By definition, $\mathcal{D}\left(s, s'\right) = \mathcal{D}\left(s', s\right)$.

For scenarios $s, s' \in \mathcal{S}$, parameter $\mathbf{t}\left(s, s'\right)$ represents the latest time period t such that realizations of all exogenous parameters resolved up till and including period t are the same in scenarios s, s'. In other words, $\mathbf{t}\left(s, s'\right)$ is the last time period at the end of which scenarios s, s' are indistinguishable based on the resolution of exogenous uncertainty. Mathematically,

$$\mathbf{t}\left(s, s'\right) = \max_t \left\{t | t \in \mathcal{T},\ \xi_\tau^s = \xi_\tau^{s'}\ \forall \tau \in \mathcal{T}, \tau \leq t\right\}$$

If $\left\{t | t \in \mathcal{T}, \xi_\tau^s = \xi_\tau^{s'}\ \forall \tau \in \mathcal{T}, \tau \leq t\right\} = \emptyset$, then we define $\mathbf{t}\left(s, s'\right) = 0$. Note that there cannot be distinct scenarios $s, s' \in \mathcal{S}$ such that $|\mathcal{D}\left(s, s'\right)| = 0$ and $\mathbf{t}\left(s, s'\right) = T$. This is because if s, s' satisfied the above conditions then they would be completely identical. By definition, $\mathbf{t}\left(s, s'\right) = \mathbf{t}\left(s', s\right)$.

$\mathcal{L}^0 = \left\{\left(s, s'\right) | s, s' \in \mathcal{S}, s < s', |\mathcal{D}\left(s, s'\right)| = 0\right\}$ represents the set of scenario pairs $\left(s, s'\right)$ such that scenarios s and s' are identical in terms of realizations for all endogenous parameters. The condition $s < s'$ prevents duplicate entries in $\mathcal{L}^0$ for the same pair of scenarios s, s'.

$\mathcal{L}^{1+} = \left\{\left(s, s'\right) | s, s' \in \mathcal{S}, s < s', |\mathcal{D}\left(s, s'\right)| \geq 1\right\}$ represents the set of scenario pairs $\left(s, s'\right)$ such that s, s' differ in realizations of θ_i for at least one $i \in \mathcal{I}$. Also, $\mathcal{L}^1 = \left\{\left(s, s'\right) | s, s' \in \mathcal{S}, s < s', |\mathcal{D}\left(s, s'\right)| = 1\right\}$.

$\mathcal{L}_T^1 = \left\{\left(s, s'\right) | \left(s, s'\right) \in \mathcal{L}^1, \mathbf{t}\left(s, s'\right) = T\right\}$ is the set of scenario pairs $\left(s, s'\right)$ such that scenarios s, s' differ in the realization of only one endogenous parameter and are identical in realizations for all exogenous parameters.

5. Model

The declarative form of stochastic programs for the class of problems described in section 3 is given by the hybrid mixed-integer disjunctive programming model (see Grossmann [7]), $(P1)$.

$(P1)$

$$\phi = \min \sum_{s \in \mathcal{S}} p^s \sum_{t \in \mathcal{T}} \left({}^w c_t^s w_t^s + {}^x c_t^s x_t^s + {}^y c_t^s y_t^s + \sum_{i \in \mathcal{I}} {}^b c_{i,t}^s b_{i,t}^s\right) \tag{2}$$

$$\text{s.t.} \sum_{\substack{\tau \in \mathcal{T}, \\ \tau \leq t}} \left({}^w A_{\tau,t}^s w_\tau^s + {}^x A_{\tau,t}^s x_\tau^s + {}^y A_{\tau,t}^s y_\tau^s + \sum_{i \in \mathcal{I}} {}^b A_{i,\tau,t}^s b_{i,\tau}^s\right) \leq a_t^s$$

$$\forall s \in \mathcal{S}, t \in \mathcal{T} \tag{3}$$

$$b_{i,1}^{s} = b_{i,1}^{s'} \quad \forall s,\ \ s' \in \mathcal{S},\ \ s < s',\ \ i \in \mathcal{I} \tag{4a}$$

$$y_1^{s} = y_1^{s'} \quad \forall s, s' \in \mathcal{S},\ \ s < s' \tag{4b}$$

$$x_t^{s} \quad = x_t^{s'} \qquad \forall (s, s') \in \mathcal{L}^0,\ \ t \in \mathcal{T},\ \ t \leq \mathbf{t}(s, s') \tag{5a}$$

$$b_{i,t+1}^{s} = b_{i,t+1}^{s'} \quad \forall (s, s') \in \mathcal{L}^0,\ \ t \in \mathcal{T},\ \ t \leq \mathbf{t}(s, s'),\ \ i \in \mathcal{I} \tag{5b}$$

$$y_{t+1}^{s} \ = y_{t+1}^{s'} \quad \forall (s, s') \in \mathcal{L}^0,\ \ t \in \mathcal{T},\ \ t \leq \mathbf{t}(s, s') \tag{5c}$$

$$\begin{bmatrix} & Z_t^{s,s'} & & \\ x_t^s & = & x_t^{s'} & \\ b_{i,t+1}^s & = & b_{i,t+1}^{s'} & \forall i \in \mathcal{I} \text{ if } t \leq T-1 \\ y_{t+1}^s & = & y_{t+1}^{s'} & \text{if } t \leq T-1 \end{bmatrix} \vee \left[\neg Z_t^{s,s'} \right]$$
$$\forall (s, s') \in \mathcal{L}^{1+},\ \ t \in \mathcal{T},\ \ t \leq \mathbf{t}(s, s') \tag{6}$$

$$Z_t^{s,s'} \Leftrightarrow \bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg b_{i,\tau}^{s} \right) \right] \ \forall \left(s, s'\right) \in \mathcal{L}^{1+},\ \ t \in \mathcal{T},\ \ t \leq \mathbf{t}\left(s, s'\right) \tag{7}$$

$$Z_t^{s,s'} \Leftrightarrow \bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg b_{i,\tau}^{s'} \right) \right] \ \forall \left(s, s'\right) \in \mathcal{L}^{1+},\ \ t \in \mathcal{T},\ \ t \leq \mathbf{t}\left(s, s'\right) \tag{8}$$

$$w_t^s \in \mathcal{W}_t^s,\ \ x_t^s \in \mathcal{X}_t^s,\ \ y_t^s \in \mathcal{Y}_t^s,\ b_{i,t}^s \in \{0, 1\} \quad \forall s \in \mathcal{S},\ \ t \in \mathcal{T},\ \ i \in \mathcal{I}$$
$$Z_t^{s,s'} \in \{True, False\}\ \ \forall \left(s, s'\right) \in \mathcal{L}^{1+},\ \ t \in \mathcal{T},\ \ t \leq \mathbf{t}\left(s, s'\right)$$

In $(P1)$, variables $b_{i,t}^s$, x_t^s and y_t^s represent the decisions to be made in time period t of scenario s. Vector w_t^s represents the other variables associated with time period t in scenario s. In process control terminology, $b_{i,t}^s$, x_t^s and y_t^s are "control variables" while w_t^s are "state variables". $b_{i,t}^s$ are binary variables while x_t^s and y_t^s are variable vectors that may have both integer and continuous components.

Parameter ${}^{b}c_{i,t}^{s}$ is the realization in scenario s for the cost coefficient corresponding to decision variable $b_{i,t}$. In other words, ${}^{b}c_{i,t}^{s} = {}^{b}c_{i,t}\left(\xi_1^s, \xi_2^s, \ldots, \xi_T^s,\ \ \theta_1^s, \theta_2^s, \ldots, \theta_I^s\right)$. Similarly, the realizations for cost coefficients corresponding to variables $x_t^{(\cdot)}$, $y_t^{(\cdot)}$ and $w_t^{(\cdot)}$ in scenario s are represented by ${}^{x}c_t^s$, ${}^{y}c_t^s$ and ${}^{w}c_t^s$, respectively. Matrices (or vectors) ${}^{b}A_{i,\tau,t}^{s}$, ${}^{x}A_{\tau,t}^{s}$, ${}^{y}A_{\tau,t}^{s}$ and ${}^{w}A_{\tau,t}^{s}$ represent the realizations for the constraint coefficient matrices (or vectors) of these variables in scenario s. Equation (2) represents the objective of minimizing the expectation of an economic criterion. For a particular scenario, inequality (3) represents constraints that govern decisions in time period t and those that link decisions across time periods. These include the square system of equality constraints which can be used to eliminate "state" variables w_t^s.

Decisions for different scenarios are linked by non-anticipativity constraints (4)-(8). The non-anticipativity rule requires that if scenarios s and s' are indistinguishable at some time, then decisions in scenarios s and s' should be the same at that time. Based on the sequence of events described in section 3, uncertainty is resolved in time period t after the implementation of decisions y_t^s and $b_{i,t}^s$. Thus, decisions $x_t^{(\cdot)}$, $b_{i,t+1}^{(\cdot)}$ and $y_{t+1}^{(\cdot)}$ should be the same for scenarios s, s' if scenarios s, s' are indistinguishable after the resolution of uncertainty in time period t. Note that in this paper, we refer to the "indistinguishability of two scenarios *after* the resolution of exogenous and endogenous uncertainty in time period t" simply by the "indistinguishability of the two scenarios *in* time period t".

Based on the sequence of events in each time period, all scenarios are indistinguishable before decisions $b^s_{i,t}$ and y^s_t are implemented in the first time period. Thus, decisions $b^{(\cdot)}_{i,1}$ and $y^{(\cdot)}_1$ have to be the same for all scenarios (equality constraints (4)). Note that the condition $s < s'$ is imposed to avoid duplication of constraints (4) for the same pair of scenarios s, s'.

Equality constraints (5) represent non-anticipativity constraints linking scenarios s, s' such that $(s, s') \in \mathcal{L}^0$; *i.e.*, the realizations of all endogenous parameters in scenarios s and s' are identical. In this case, scenarios s, s' will be indistinguishable in time period t if and only if these scenarios are identical in realizations of all exogenous parameters observed up till and including time period t. Accordingly, constraint (5) applies non-anticipativity constraints on decisions $x^{(\cdot)}_t$, $y^{(\cdot)}_{t+1}$, $b^{(\cdot)}_{i,t+1}$ for scenarios s, s' only if t satisfies $t \leq \mathbf{t}(s, s')$.

Constraints (6)–(8) are non-anticipativity constraints linking scenarios s, s' such that $(s, s') \in \mathcal{L}^{1+}$; *i.e.*, scenarios s and s' differ in the realization of at least one endogenous parameter. In this case, the indistinguishability of scenarios s, s' in time period t depends on both, endogenous and exogenous uncertainty resolved in the past. Boolean variable $Z^{s,s'}_t$ is $True$ if and only if scenarios s and s' are indistinguishable after the resolution of uncertainty in time period t. Clearly, for $t > \mathbf{t}(s, s')$ scenarios s, s' can be distinguished simply based on realizations of the exogenous parameters. Hence, $Z^{s,s'}_t = False$ for $t > \mathbf{t}(s, s')$. Therefore, constraints (6)–(8) are applied only for t such that $t \leq \mathbf{t}(s, s')$, where $(s, s') \in \mathcal{L}^{1+}$.

Disjunction (6) imposes the non-anticipativity constraints on variables $x^{(\cdot)}_t$, $y^{(\cdot)}_{t+1}$, $b^{(\cdot)}_{i,t+1}$ for scenarios s, s' only if $Z^{s,s'}_t$ is $True$, *i.e.*, if scenarios s and s' are indistinguishable in time period t. By definition of $\mathbf{t}(s, s')$, if $t \leq \mathbf{t}(s, s')$, the indistinguishability of scenarios s, s' in time period t depends purely on the endogenous uncertainty resolved through the decisions. Logic constraints (7) and (8) relate the indistinguishability of scenarios s, s' in time period t with decisions $b^s_{i,\tau}$ and $b^{s'}_{i,\tau}$ respectively. Scenarios s, s' differ in realizations for a finite set of endogenous parameters. Constraint (7) states[1] that $Z^{s,s'}_t$ is $True$ if and only if uncertainty has not been resolved in any of these parameters up till and including time period t of scenario s. Similarly, constraint (8) relates variables $Z^{s,s'}_t$ to the corresponding decision variables for scenario s'.

Note that to account for the offset in the time index of these variables, the non-anticipativity constraints on variables $b^{(\cdot)}_{i,t+1}$, $y^{(\cdot)}_{t+1}$ for scenarios s, s' are applied only if $t \leq T - 1$. Although it may seem that a similar restriction is needed in constraints (5b)–(5c), however, as explained earlier in this section, we cannot have distinct scenarios

[1] In theory, the logical operator "$\neg$" should only be used with Boolean variables. Since $b^s_{i,t}$ are binary variables, therefore constraints (7) and (8) involve a slight inconsistency in notation. A more rigorous formulation can be obtained at the expense of additional notation by defining constraints (7) and (8) in terms of Boolean variables $B^s_{i,t}$ and specifying an equivalence between variables $B^s_{i,t}$ and $b^s_{i,t}$.

$s, s' \in \mathcal{S}$ such that $(s, s') \in \mathcal{L}^0$ and $\mathbf{t}(s, s') = T$. Hence, in constraint (5b)–(5c), the condition that $t \leq T - 1$ is implicit in the condition $t \leq \mathbf{t}(s, s')$.

$\mathcal{W}_t^s$, $\mathcal{X}_t^s$ and $\mathcal{Y}_t^s$ represent the bounds and integrality restrictions on variables w_t^s, x_t^s and y_t^s respectively, for all $t \in \mathcal{T}$, $s \in \mathcal{S}$.

6. Model properties

In this section, we present a set of properties that lead to reduction in the dimensionality of model $(P1)$. We show that constraint (8) is redundant and that instead of applying constraints (6) and (7) for all pairs of scenarios (s, s') such that s, s' differ in the realization of at least one endogenous parameter, these constraints need to be applied for s, s' only if these scenarios differ in the realization of exactly one endogenous parameter and are identical in realizations for all exogenous parameters. Note that we will use b^s to represent the vector of variables $b_{i,t}^s$ for all (i, t). Similarly, vector b will represent the vector of b^s for all s. The same convention will be used to represent vectors of variables w_t^s, x_t^s, y_t^s, $Z_t^{s,s'}$ and parameters that are introduced later in the paper. The tuple (b, w, x, y, Z) will be used to represent a solution to the model under consideration. Further, in all properties presented in this paper, it is assumed that variables $b_{i,t}^s \in \{0, 1\}$ and $Z_t^{s,s'} \in \{True, False\}$. Similarly, solutions $\hat{b}_{i,t}^s \in \{0, 1\}$ and $\hat{Z}_t^{s,s'} \in \{True, False\}$.

Proposition 1. *Consider constraints (9)–(11) for given $s, s', \hat{t}$ where $s, s' \in \mathcal{S}, \hat{t} \in \mathcal{T}, \hat{t} \leq T - 1$.*

$$b_{i,1}^s = b_{i,1}^{s'} \ \forall i \in \mathcal{I} \tag{9}$$

$$\begin{bmatrix} & Z_t^{s,s'} & \\ b_{i,t+1}^s & = & b_{i,t+1}^{s'} \forall i \in \mathcal{I} \end{bmatrix} \vee \left[\neg Z_t^{s,s'}\right] \ \forall t \in \mathcal{T}, \ t \leq \hat{t} \tag{10}$$

$$Z_t^{s,s'} \Leftrightarrow \bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg b_{i,\tau}^s\right)\right] \ \forall t \in \mathcal{T}, \ t \leq \hat{t} \tag{11}$$

If vectors $\hat{b}^s$, $\hat{b}^{s'}$, $\hat{Z}^{s,s'}$ satisfy constraints (9)–(11),

(a) For $t \in \mathcal{T}, t \leq \hat{t}$,

$$\bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}_{i,\tau}^s\right)\right] \Rightarrow \left[\hat{b}_{i,\tau}^s = \hat{b}_{i,\tau}^{s'} \ \forall i \in \mathcal{I}, \ \tau \in \mathcal{T}, \ \tau \leq t+1\right]$$

(b) For $t = \hat{t} + 1$,

$$\bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}_{i,\tau}^s\right)\right] \Rightarrow \left[\hat{b}_{i,\tau}^s = \hat{b}_{i,\tau}^{s'} \ \forall i \in \mathcal{I}, \ \tau \in \mathcal{T}, \ \tau \leq t\right]$$

(c) For $t \in \mathcal{T}, t \leq \hat{t}+1$,

$$\bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}_{i,\tau}^{s} \right) \right] \Leftrightarrow \bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}_{i,\tau}^{s'} \right) \right]$$

(d) For $t \in \mathcal{T}, t \leq \hat{t}$,

$$\bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}_{i,\tau}^{s'} \right) \right] \Rightarrow \left[\hat{b}_{i,\tau}^{s} = \hat{b}_{i,\tau}^{s'} \ \forall i \in \mathcal{I}, \ \ \tau \in \mathcal{T}, \ \ \tau \leq t+1 \right]$$

(e) For $t = \hat{t}+1$,

$$\bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}_{i,\tau}^{s'} \right) \right] \Rightarrow \left[\hat{b}_{i,\tau}^{s} = \hat{b}_{i,\tau}^{s'} \ \forall i \in \mathcal{I}, \ \ \tau \in \mathcal{T}, \ \ \tau \leq t \right]$$

Note that the left hand sides of (a) and (b) involve variables $b_{i,\tau}^{s}$ while the left hand sides of (d) and (e) involve variables $b_{i,\tau}^{s'}$.

Proof. See Appendix A. □

We use Proposition 1 as a basis to prove the following theorem.

Theorem 1. *If solution $\left(\hat{b}, \hat{w}, \hat{x}, \hat{y}, \hat{Z}\right)$ satisfies constraints (4a), (6) and (7), then it also satisfies constraint (8). Thus, constraint (8) is redundant in $(P1)$.*

Proof. Suppose solution $\left(\hat{b}, \hat{w}, \hat{x}, \hat{y}, \hat{Z}\right)$ satisfies constraints (4a), (6) and (7). Consider scenarios $s_a, s_b \in \mathcal{S}$ such that $(s_a, s_b) \in \mathcal{L}^{1+}$. We will prove that solution $\left(\hat{b}, \hat{w}, \hat{x}, \hat{y}, \hat{Z}\right)$ satisfies constraint (8) for $\left(s, s'\right) = (s_a, s_b)$. The theorem follows as a result.

By definition, $\mathbf{t}\,(s_a, s_b) = \max_t \left\{ t | t \in \mathcal{T}, \xi_\tau^{s_a} = \xi_\tau^{s_b} \ \forall \tau \in \mathcal{T}, \tau \leq t \right\}$. If $\{t | t \in \mathcal{T}, \xi_\tau^{s_a} = \xi_\tau^{s_b} \ \forall \tau \in \mathcal{T}, \tau \leq t \} = \emptyset$, then by convention $\mathbf{t}\,(s_a, s_b) = 0$. Hence, $\left(\hat{b}, \hat{w}, \hat{x}, \hat{y}, \hat{Z}\right)$ satisfies constraint (8) vacuously for $\left(s, s'\right) = (s_a, s_b)$.

On the other hand, $\mathbf{t}\,(s_a, s_b) \geq 1$ if $\left\{ t | t \in \mathcal{T}, \xi_\tau^{s_a} = \xi_\tau^{s_b} \ \forall \tau \in \mathcal{T}, \tau \leq t \right\} \neq \emptyset$. Since $\left(\hat{b}, \hat{w}, \hat{x}, \hat{y}, \hat{Z}\right)$ satisfies constraints (4a), (6) and (7) and $(s_a, s_b) \in \mathcal{L}^{1+}$, therefore subvectors $\hat{b}^{s_a}, \hat{b}^{s_b}, \hat{Z}^{s_a, s_b}$ satisfy constraints (9)–(11) for $\left(s, s'\right) = (s_a, s_b), \hat{t} = \min\,(T-1, \mathbf{t}\,(s_a, s_b))$. (The equality constraint on variables $b_{i,t+1}^{(\cdot)}$ inside disjunction (6) is applied only if $t \leq T-1$. Hence, $\left(\hat{b}^{s_a}, \hat{b}^{s_b}, \hat{Z}^{s_a, s_b}\right)$ is guaranteed to satisfy disjunction (10) only for $t \leq \min\,(T-1, \mathbf{t}\,(s_a, s_b))$).

Using result (c) of Proposition 1, we get

$$\bigwedge_{i \in \mathcal{D}(s_a,s_b)} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}_{i,\tau}^{s_a} \right) \right] \Leftrightarrow \bigwedge_{i \in \mathcal{D}(s_a,s_b)} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}_{i,\tau}^{s_b} \right) \right] \quad \forall t \in \mathcal{T}, \ \ t \leq \min\,(T-1, \mathbf{t}\,(s_a, s_b)) + 1$$

Since $\mathbf{t}(s_a, s_b) \leq T$, therefore $\min(T-1, \mathbf{t}(s_a, s_b))+1 = \min(T, \mathbf{t}(s_a, s_b)+1) \geq \mathbf{t}(s_a, s_b)$. Hence,

$$\bigwedge_{i \in \mathcal{D}(s_a, s_b)} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}_{i,\tau}^{s_a} \right) \right] \Leftrightarrow \bigwedge_{i \in \mathcal{D}(s_a, s_b)} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}_{i,\tau}^{s_b} \right) \right] \quad \forall t \in \mathcal{T}, \quad t \leq \mathbf{t}(s_a, s_b) \tag{12}$$

Since sub-vectors $\hat{b}^{s_a}, \hat{b}^{s_b}, \hat{Z}^{s_a, s_b}$ satisfy logic constraint (7) for $(s, s') = (s_a, s_b)$, we can combine constraint (7) with equation (12) to infer that sub-vectors $\hat{b}^{s_a}, \hat{b}^{s_b}, \hat{Z}^{s_a, s_b}$ satisfy logic constraint (8). The result follows. □

Proposition 2. *Consider constraints* (13)-(16) *in variables* b, x, y, Z *defined over the tuple* (s, s', t)

$$b_{i,t}^{s} = b_{i,t}^{s'} \quad \forall i \in \mathcal{I} \tag{13a}$$

$$y_t^s = y_t^{s'} \tag{13b}$$

$$x_t^s = x_t^{s'} \tag{14a}$$

$$b_{i,t+1}^{s} = b_{i,t+1}^{s'} \quad \forall i \in \mathcal{I} \tag{14b}$$

$$y_{t+1}^s = y_{t+1}^{s'} \tag{14c}$$

$$\begin{bmatrix} & Z_t^{s,s'} & & & \\ x_t^s & = & x_t^{s'} & & \\ b_{i,t+1}^s & = & b_{i,t+1}^{s'} & \forall i \in \mathcal{I} & \text{if } t \leq T-1 \\ y_{t+1}^s & = & y_{t+1}^{s'} & & \text{if } t \leq T-1 \end{bmatrix} \vee \left[\neg Z_t^{s,s'} \right] \tag{15}$$

$$Z_t^{s,s'} \Leftrightarrow \bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg b_{i,\tau}^{s} \right) \right] \tag{16}$$

Suppose the set of scenarios is equivalent to $\Xi \times (\times_{i \in \mathcal{I}} \Theta_i)$ *and vectors* $\hat{b}, \hat{x}, \hat{y}, \hat{Z}$ *satisfy*

(i) Constraints (13a)–(13b) for (s, s', t) *such that* $s, s' \in \mathcal{S}, s < s', t = 1$
(ii) Constraint (14) for (s, s', t) *such that* $(s, s') \in \mathcal{L}^0, t \in \mathcal{T}, t \leq \boldsymbol{t}(s, s')$
(iii) Constraints (15)–(16) for (s, s', t) *such that* $(s, s') \in \mathcal{L}_T^1, t \in \mathcal{T}$

Then, $\hat{b}, \hat{x}, \hat{y}, \hat{Z}$ *satisfy constraints (15)–(16) for* (s, s', t) *such that* $(s, s') \in \mathcal{L}^{1+}, t \in \mathcal{T}, t \leq \boldsymbol{t}(s, s')$.

Proof. See Appendix C (Based on Lemma 1, Appendix B). □

Based on the above proposition, we define model $(P2)$ where constraints (17) and (18) are applied instead of constraints (6) and (7), respectively. Also, logic constraint (8) has been dropped.

$(P2)$

$$\phi = \min \sum_{s\in\mathcal{S}} p^s \sum_{t\in\mathcal{T}} \left({}^w c_t^s w_t^s + {}^x c_t^s x_t^s + {}^y c_t^s y_t^s + \sum_{i\in\mathcal{I}} {}^b c_{i,t}^s b_{i,t}^s \right)$$

s.t. (3), (4), (5)

$$\begin{bmatrix} & Z_t^{s,s'} & & \\ x_t^s & = & x_t^{s'} & \\ b_{i,t+1}^s & = & b_{i,t+1}^{s'} & \forall i \in \mathcal{I} \text{ if } t \le T-1 \\ y_{t+1}^s & = & y_{t+1}^{s'} & \text{if } t \le T-1 \end{bmatrix} \vee \left[\neg Z_t^{s,s'}\right]$$

$$\forall \left(s, s'\right) \in \mathcal{L}_T^1, t \in \mathcal{T} \quad (17)$$

$$Z_t^{s,s'} \Leftrightarrow \left[\bigwedge_{\tau=1}^{t} \left(\neg b_{i,\tau}^s\right)\right] \qquad \forall \left(s, s'\right) \in \mathcal{L}_T^1, t \in \mathcal{T}, \{i\} = \mathcal{D}\left(s, s'\right) \quad (18)$$

$$w_t^s \in \mathcal{W}_t^s,\ x_t^s \in \mathcal{X}_t^s,\ y_t^s \in \mathcal{Y}_t^s,\ b_{i,t}^s \in \{0,1\} \quad \forall s \in \mathcal{S}, t \in \mathcal{T}, i \in \mathcal{I}$$

$$Z_t^{s,s'} \in \{True, False\} \qquad \forall \left(s, s'\right) \in \mathcal{L}_T^1, t \in \mathcal{T}$$

Theorem 2. *If $\left(\hat{b}, \hat{w}, \hat{x}, \hat{y}, \hat{Z}\right)$ is an optimal solution of $(P1)$ then it is also an optimal solution of $(P2)$, and vice versa.*

Proof. Since the objective functions of $(P1)$ and $(P2)$ are the same, it is sufficient to show that the feasible regions of $(P1)$ and $(P2)$ are the same.

Suppose $\left(\hat{b}, \hat{w}, \hat{x}, \hat{y}, \hat{Z}\right)$ is a feasible solution of $(P1)$. Compare models $(P1)$ and $(P2)$. Constraints (3)–(5) are common to both models while disjunctions (17) and (6) differ only in the domain for $\left(s, s', t\right)$. Constraint (18) differs from (7) in the domain for $\left(s, s', t\right)$ and in the right hand side of the logic relationship.

Let $\mathcal{F}_1$ denote the domain of $\left(s, s', t\right)$ in constraints (6)–(7) and let $\mathcal{F}_2$ denote the domain of $\left(s, s', t\right)$ in constraints (17)–(18). Thus,

$$\mathcal{F}_1 = \left\{\left(s, s', t\right) \mid \left(s, s'\right) \in \mathcal{L}^{1+}, t \in \mathcal{T}, t \le \mathbf{t}\left(s, s'\right)\right\}$$

$$\mathcal{F}_2 = \left\{\left(s, s', t\right) \mid \left(s, s'\right) \in \mathcal{L}_T^1, t \in \mathcal{T}\right\}.$$

where,

$$\mathcal{L}_T^1 = \left\{\left(s, s'\right) \mid \left(s, s'\right) \in \mathcal{L}^1, \mathbf{t}\left(s, s'\right) = T\right\}$$

By definition,

$$\mathcal{L}_T^1 \subseteq \mathcal{L}^1 \subseteq \mathcal{L}^{1+}.$$

Now,

$$\begin{aligned} \mathcal{F}_2 &= \left\{\left(s, s', t\right) \mid \left(s, s'\right) \in \mathcal{L}_T^1, t \in \mathcal{T}\right\} \\ &\equiv \left\{\left(s, s', t\right) \mid \left(s, s'\right) \in \mathcal{L}_T^1, t \in \mathcal{T}, t \le \mathbf{t}\left(s, s'\right)\right\} \quad \left(\text{because } \mathbf{t}\left(\mathrm{s}, \mathrm{s}'\right) = \mathrm{T} \text{ for } \left(\mathrm{s}, \mathrm{s}'\right) \in \mathcal{L}_\mathrm{T}^1\right) \\ &\subseteq \left\{\left(s, s', t\right) \mid \left(s, s'\right) \in \mathcal{L}^{1+}, t \in \mathcal{T}, t \le \mathbf{t}\left(s, s'\right)\right\} \; \left(\text{because } \mathcal{L}_\mathrm{T}^1 \subseteq \mathcal{L}^{1+}\right) \\ &= \mathcal{F}_1 \end{aligned}$$

Also, $|\mathcal{D}(s, s')| = 1$ for $(s, s', t) \in \mathcal{F}_2$. Therefore, the right hand side of constraint (7) reduces to the right hand side of constraint (18). Since $\mathcal{F}_2 \subseteq \mathcal{F}_1$, therefore $(P2)$ is a relaxation of $(P1)$. Thus $\left(\hat{b}, \hat{w}, \hat{x}, \hat{y}, \hat{Z}\right)$ should be a feasible solution of $(P2)$.

Conversely, suppose $\left(\hat{b}, \hat{w}, \hat{x}, \hat{y}, \hat{Z}\right)$ is a feasible solution of $(P2)$. Thus, $\left(\hat{b}, \hat{w}, \hat{x}, \hat{y}, \hat{Z}\right)$ satisfies constraints (3), (4), (5), (17) and (18). Using Proposition 2, we can infer that $\left(\hat{b}, \hat{w}, \hat{x}, \hat{y}, \hat{Z}\right)$ satisfies constraints (6) and (7). Further, using Theorem 1 we can infer that $\left(\hat{b}, \hat{w}, \hat{x}, \hat{y}, \hat{Z}\right)$ satisfies logic constraint (8). Thus, $\left(\hat{b}, \hat{w}, \hat{x}, \hat{y}, \hat{Z}\right)$ is a feasible solution of $(P1)$. □

The following remarks can be made about the proposed model.

1. The "standard" stochastic programming formulation (SSP) is clearly a specific case of model $(P2)$ when the problem involves exogenous uncertainty only $\left(\mathcal{L}^{1+} = \mathcal{L}_T^1 = \emptyset\right)$.
2. The proofs of Proposition 1, Theorem 1 and Theorem 2 are independent of the choice of the set of scenarios. To illustrate the dependence of Proposition 2 on the set of scenarios, consider indices $s_a, s_b \in \mathcal{S}$ such that the corresponding scenarios differ in the realizations of r endogenous parameters, where $r = |\mathcal{D}(s_a, s_b)| \geq 1$. Broadly, Proposition 2, which is used in Theorem 2, is based on the assumption that there exist indices $s_1, s_2, \ldots, s_r \in \mathcal{S}$ such that $(s_a, s_1), (s_1, s_2), (s_2, s_3), \ldots, (s_{r-1}, s_r) \in \mathcal{L}_T^1$ while $(s_r, s_b) \in \mathcal{L}^0$. Proposition 2 is then a result of the fact that the non-anticipativity constraints linking s_a with s_b are implied by the "chaining" of non-anticipativity constraints linking s_a with s_1, s_1 with s_2, s_2 with $s_3, \ldots, s_{r-1}$ with s_r and s_r with s_b. Since we choose the set of scenarios as $\Xi \times (\times_{i\in\mathcal{I}}\Theta_i)$, for any realization of the vector of exogenous parameters ξ, the set of scenarios includes all possible combinations of realizations for the endogenous parameters. Thus, we can generate r "intermediate" scenarios from scenario s_a by progressively changing the realization of one of the r distinguishing endogenous parameters to the corresponding realization in scenario s_b. The realizations of all exogenous parameters in these r scenarios are identical to those in s_a. Since $s_a \in \mathcal{S}$, these r scenarios also belong to the set of scenarios. Thus, we can choose indices $s_1, s_2, \ldots, s_r \in \mathcal{S}$ for these r scenarios. Hence, the non-anticipativity constraints for s_a, s_b follow by "chaining", as explained above.
3. The models and proofs presented here are based on the assumption that the endogenous uncertainty associated with source i can be represented by one parameter. Thus, θ_i is a scalar. To consider the more general case, suppose θ_i is an $n_i \times 1$ vector. For example, in the gas field problem, the uncertainty in a field is represented by uncertainty in the size and quality of the field. Therefore, in that problem $n_i = 2$ for each field i.
   If we choose the set of scenarios as $\Xi \times (\times_{i\in\mathcal{I}}\Theta_i)$, where $(\times_{i\in\mathcal{I}}\Theta_i)$ represents all possible combinations of realizations for *vectors* θ_i for all i, we can again use the "chaining" argument to prove that the solutions to models $(P1)$ and $(P2)$ are the same for

$$\mathcal{L}_T^1 = \left\{\left(s, s'\right) \mid s, s' \in \mathcal{S}, s < s', \mathbf{t}\left(s, s'\right) = T,\right.$$

$$\exists\left(i^*, l^*\right), i^* \in \mathcal{I}, l^* \in \{1, 2, \ldots, n_{i^*}\} \text{ such that}$$
$$\theta^s_{l^*,i^*} \neq \theta^{s'}_{l^*,i^*},$$
$$\theta^s_{l,i} = \theta^{s'}_{l,i}\ \forall l \in \{1, 2, \ldots, n_i\}, i \in \mathcal{I} \setminus \left\{i^*\right\}\Big\}$$

However, stronger results may be obtained if the set of scenarios is chosen as $\Xi \times \left(\times_{i\in\mathcal{I}}\left(\times_{l=1}^{n_i}\Theta_{l,i}\right)\right)$, where $\Theta_{l,i}$ represents the set of possible realizations for endogenous uncertain parameter $\theta_{l,i}$ associated with source i. The "chaining" argument can then be used to prove that the solutions to models $(P1)$ and $(P2)$ are the same for

$$\begin{aligned}\mathcal{L}^1_T = \Big\{\left(s, s'\right) \mid\ & s, s' \in \mathcal{S}, s < s', \mathbf{t}\left(s, s'\right) = T,\\ & \exists\left(i^*, l^*\right), i^* \in \mathcal{I}, l^* \in \{1, 2, \ldots, n_{i^*}\} \text{ such that}\\ & \theta^s_{l^*,i^*} \neq \theta^{s'}_{l^*,i^*},\\ & \theta^s_{l,i^*} = \theta^{s'}_{l,i^*}\ \forall l \in \{1, 2, \ldots, n_{i^*}\} \setminus \{l^*\},\\ & \theta^s_{l,i} = \theta^{s'}_{l,i}\ \forall l \in \{1, 2, \ldots, n_i\}, i \in \mathcal{I} \setminus \{i^*\}\Big\}\end{aligned}$$

In other words, the disjunctive non-anticipativity constraints will need to be applied between scenarios s, s' only if the two scenarios differ in the realization of exactly one endogenous *scalar* parameter, θ_{l^*,i^*}.

4. In the present form, $(P2)$ has disjunctions and linear constraints linking Boolean, binary and continuous variables. The model can be reformulated as a mixed integer linear program by representing Boolean variables $Z^{s,s'}_t$ as 0-1 variables $z^{s,s'}_t$, and reformulating the logic constraints and disjunctions as linear constraints by using big-M or convex hull reformulations (see Balas [2], Turkay and Grossmann [21]). Specifically, logic constraint (18) can be replaced by linear constraints (19)-(20).

$$z^{s,s'}_t \leq 1 - b^s_{i,\tau} \qquad \forall\left(s, s'\right) \in \mathcal{L}^1_T; t, \tau \in \mathcal{T}, \tau \leq t, \{i\} = \mathcal{D}\left(s, s'\right) \quad (19)$$

$$z^{s,s'}_t \geq 1 - \sum_{\tau\leq t, \tau\in\mathcal{T}} b^s_{i,\tau} \quad \forall\left(s, s'\right) \in \mathcal{L}^1_T, t \in \mathcal{T}, \{i\} = \mathcal{D}\left(s, s'\right) \quad (20)$$

It should be noted that if constraints (19)–(20) are used as a reformulation of constraint (18), variables $z^{s,s'}_t$ will take 0–1 values even if they are defined as continuous variables with bounds $0 \leq z^{s,s'}_t \leq 1$. This follows because $b^s_{i,\tau} \in \{0, 1\}$ for all $s \in \mathcal{S}, t \in \mathcal{T}, i \in \mathcal{I}$.

In spite of the above property, the MILP reformulation of $(P2)$ can easily become very large for real-world problems, making solution with a generic MILP solver highly inefficient. In the next section, we present a specialized branch and bound algorithm motivated by the work of Caroe and Schultz [4].

7. Branch and bound algorithm

Model $(P2)$ is coupled in scenarios through the non-anticipativity constraints. In the proposed branch and bound algorithm, lower bounds at each node are generated by solving a Lagrangean dual problem which is obtained by relaxing disjunction (17) and replacing equality constraints (4) and (5) by penalty terms in the objective. Each sub-problem in the Lagrangean dual problem corresponds to an MILP for one of the scenarios. We first illustrate the formulation of the Lagrangean dual for the problem at the root node.

$$\left(P2_{ref}\right)$$

$$\begin{aligned}
\phi_{ref} = \; & \min \sum_{s\in\mathcal{S}} p^s \sum_{t\in\mathcal{T}} \left({}^w c_t^s w_t^s + {}^x c_t^s x_t^s + {}^y c_t^s y_t^s + \sum_{i\in\mathcal{I}} {}^b c_{i,t}^s b_{i,t}^s \right) \\
\text{s.t.} \quad & (3), (4), (5), (17), (18) \\
& Z_t^{s',s} \Leftrightarrow \left[\bigwedge_{\tau=1}^{t} \left(\neg b_{i,\tau}^{s'}\right) \right] \quad \forall \left(s, s'\right) \in \mathcal{L}_T^1, t \in \mathcal{T}, \{i\} = \mathcal{D}\left(s, s'\right) \qquad (21) \\
& Z_t^{s',s} = Z_t^{s,s'} \qquad \forall \left(s, s'\right) \in \mathcal{L}_T^1, t \in \mathcal{T} \qquad (22) \\
& w_t^s \in \mathcal{W}_t^s,\; x_t^s \in \mathcal{X}_t^s,\; y_t^s \in \mathcal{Y}_t^s,\; b_{i,t}^s \in \{0,1\} \; \forall s \in \mathcal{S}, t \in \mathcal{T}, i \in \mathcal{I} \\
& Z_t^{s,s'}, Z_t^{s',s} \in \{True, False\} \qquad \forall \left(s, s'\right) \in \mathcal{L}_T^1, t \in \mathcal{T}
\end{aligned}$$

Model $\left(P2_{ref}\right)$ is a reformulation of $(P2)$ obtained by introducing dummy variables $Z_t^{s',s}$ for $(s, s') \in \mathcal{L}_T^1$, $t \in \mathcal{T}$. In $(P2_{ref})$, constraint (21) relates variables $Z_t^{s',s}$ to variables $b^{s'}$ while constraint (18) relates variables $Z_t^{s,s'}$ to variables b^s. Equality constraint (22) represents the symmetry restriction on $Z_t^{s,s'}$ with respect to s, s'. From the proofs presented in the previous section, it is clear that any solution of model $(P2)$ will also satisfy constraints (21) and (22). Therefore, the optimal solutions of $(P2)$ and $(P2_{ref})$ will be the same and $\phi = \phi_{ref}$. While constraints (21) and (22) are redundant (as shown in Theorem 1), inclusion of these constraints in the model tightens the Lagrangean dual and hence the lower bounds. A qualitative reasoning for the same will be presented later in this section.

Model $(P2_{RLR})$ is obtained from $(P2_{ref})$ by relaxing disjunctions (17) and replacing equality constraints (4), (5) and (22) by penalty terms in the objective.

$$\begin{aligned}
(P2_{RLR}) \quad & \phi_{RLR}({}^b\lambda, {}^x\lambda, {}^y\lambda, {}^z\lambda) \\
= \min \; & \sum_{s\in\mathcal{S}} p^s \sum_{t\in\mathcal{T}} \left({}^w c_t^s w_t^s + {}^x c_t^s x_t^s + {}^y c_t^s y_t^s + \sum_{i\in\mathcal{I}} {}^b c_{i,t}^s b_{i,t}^s \right) \\
& + \sum_{\substack{s,s'\in\mathcal{S} \\ s<s'}} \left[\sum_{i\in\mathcal{I}} {}^b\lambda_{i,0}^{s,s'} \left(b_{i,1}^s - b_{i,1}^{s'}\right) + {}^y\lambda_0^{s,s'} \left(y_1^s - y_1^{s'}\right) \right] \\
& + \sum_{(s,s')\in\mathcal{L}^0} \sum_{t=1}^{\mathbf{t}(s,s')} \left(\sum_{i\in\mathcal{I}} {}^b\lambda_{i,t}^{s,s'} \left(b_{i,t+1}^s - b_{i,t+1}^{s'}\right) + {}^y\lambda_t^{s,s'} \left(y_{t+1}^s - y_{t+1}^{s'}\right) \right. \\
& \qquad \left. + {}^x\lambda_t^{s,s'} \left(x_t^s - x_t^{s'}\right) \right)
\end{aligned}$$

$$+\sum_{(s,s')\in\mathcal{L}_T^1}\sum_{t\in\mathcal{T}} {}^{z}\lambda_t^{s,s'}\left(z_t^{s,s'}-z_t^{s',s}\right)$$

s.t. (3)

$$z_t^{s,s'} \le 1 - b_{i,\tau}^s \quad \forall (s,s')\in\mathcal{L}_T^1, t\in\mathcal{T}, \{i\}=\mathcal{D}(s,s'), \tau\in\mathcal{T}, \tau\le t \tag{23}$$

$$z_t^{s,s'} \ge 1 - \sum_{\substack{\tau\in\mathcal{T}\\ \tau\le t}} b_{i,\tau}^s \quad \forall (s,s')\in\mathcal{L}_T^1, t\in\mathcal{T}, \{i\}=\mathcal{D}(s,s') \tag{24}$$

$$z_t^{s',s} \le 1 - b_{i,\tau}^{s'} \quad \forall (s,s')\in\mathcal{L}_T^1, t\in\mathcal{T}, \{i\}=\mathcal{D}(s,s'), \tau\in\mathcal{T}, \tau\le t \tag{25}$$

$$z_t^{s',s} \ge 1 - \sum_{\substack{\tau\in\mathcal{T}\\ \tau\le t}} b_{i,\tau}^{s'} \quad \forall (s,s')\in\mathcal{L}_T^1, t\in\mathcal{T}, \{i\}=\mathcal{D}(s,s') \tag{26}$$

$$w_t^s \in \mathcal{W}_t^s,\ x_t^s\in\mathcal{X}_t^s,\ y_t^s\in\mathcal{Y}_t^s,\ b_{i,t}^s\in\{0,1\} \qquad \forall s\in\mathcal{S}, t\in\mathcal{T}, i\in\mathcal{I}$$

$$0\le z_t^{s,s'}, z_t^{s',s}\le 1 \qquad \forall (s,s')\in\mathcal{L}_T^1, t\in\mathcal{T}$$

The parameters ${}^{b}\lambda_{i,t}^{s,s'}$, ${}^{x}\lambda_t^{s,s'}$, ${}^{y}\lambda_t^{s,s'}$ and ${}^{z}\lambda_t^{s,s'}$ represent Lagrange multipliers corresponding to constraints $b_{i,t}^s = b_{i,t}^{s'}$, $x_t^s = x_t^{s'}$, $y_t^s = y_t^{s'}$ and $z_t^{s,s'} = z_t^{s',s}$, respectively. The Boolean variables $Z_t^{s,s'}$ have been replaced by continuous variables $z_t^{s,s'}$. As explained in section 6, variables $z_t^{s,s'}$ will take 0–1 values even if the integrality condition is not imposed on these variables. Constraints (23)–(24) and (25)–(26) are linear algebraic formulations of logic constraints (18) and (21) respectively. $(P2_{RLR})$ is an MILP model and clearly a relaxation of $(P2_{ref})$ (and hence of $(P2)$) for any values of the Lagrange multipliers. Thus,

$$\phi_{RLR}({}^{b}\lambda, {}^{x}\lambda, {}^{y}\lambda, {}^{z}\lambda) \le \phi \quad \forall({}^{b}\lambda, {}^{x}\lambda, {}^{y}\lambda, {}^{z}\lambda)$$

Further, $(P2_{RLR})$ can be decomposed into one MILP sub-problem for each scenario. Note that since the disjunctions have been completely relaxed, $(P2_{RLR})$ is not the Lagrangean relaxation of $(P2_{ref})$. Therefore, we refer to it as the "relaxed" Lagrangean relaxation of $(P2_{ref})$. Then, the relaxed Lagrangean dual problem (see Guignard and Kim [8], Nemhauser and Wolsey [15]) corresponding to $(P2_{ref})$ is

$$\phi_{RLD} = \max_{{}^{b}\lambda, {}^{x}\lambda, {}^{y}\lambda, {}^{z}\lambda} \phi_{RLR}\left({}^{b}\lambda, {}^{x}\lambda, {}^{y}\lambda, {}^{z}\lambda\right)$$

Clearly, ϕ_{RLD} gives a lower bound to ϕ. In general, the solution of the Lagrangean dual may not satisfy the relaxed disjunctions. The penalty term corresponding to equality constraint (22) tries to force $Z_t^{s,s'} = Z_t^{s',s}$ and hence $\left[\bigwedge_{\tau=1}^{t}\left(\neg b_{i,\tau}^s\right)\right] = \left[\bigwedge_{\tau=1}^{t}\left(\neg b_{i,\tau}^{s'}\right)\right]$ for all $(s,s')\in\mathcal{L}_T^1$, $t\in\mathcal{T}$, $\{i\}=\mathcal{D}(s,s')$. For $(s,s')\in\mathcal{L}_T^1$, scenarios s, s' differ only in the realization for θ_i, where $\{i\}=\mathcal{D}(s,s')$. Thus, forcing decisions for source i to follow non-anticipativity should force other decisions to follow non-anticipativity too. This motivates the inclusion of constraints (21) and (22) in model $(P2_{ref})$.

Lower bounds at each node are generated by solving one such Lagrangean dual problem. Model (P) represents the problem at any node in the branch and bound tree.

(P)

$$\phi_P = \min \sum_{s \in \mathcal{S}} p^s \sum_{t \in \mathcal{T}} \left({}^w c_t^s w_t^s + {}^x c_t^s x_t^s + {}^y c_t^s y_t^s + \sum_{i \in \mathcal{I}} {}^b c_{i,t}^s b_{i,t}^s \right)$$

$$\text{s.t.} \quad (3), (18), (21)$$

$$\left. \begin{array}{lll} x_t^s & = x_t^{s'} & \text{if } t > 0 \\ b_{i,t+1}^s & = b_{i,t+1}^{s'} \ \forall i \in \mathcal{I} & \text{if } t \le T-1 \\ y_{t+1}^s & = y_{t+1}^{s'} & \text{if } t \le T-1 \end{array} \right\} \qquad \forall (s, s', t) \in \mathcal{N}_P^e \tag{27}$$

$$\begin{bmatrix} & Z_t^{s,s'} & \\ x_t^s & = & x_t^{s'} \\ b_{i,t+1}^s & = & b_{i,t+1}^{s'} \ \forall i \in \mathcal{I} \text{ if } t \le T-1 \\ y_{t+1}^s & = & y_{t+1}^{s'} \qquad \text{if } t \le T-1 \end{bmatrix} \vee \left[\neg Z_t^{s,s'} \right]$$

$$\forall (s, s', t) \in \mathcal{N}_P^d \tag{28}$$

$$Z_t^{s,s'} = Z_t^{s',s} \qquad \forall (s, s', t) \in \mathcal{N}_P^d \tag{29}$$

$$w_t^s \in \mathcal{W}_t^s,\ x_t^s \in \mathcal{X}_{t,P}^s,\ y_t^s \in \mathcal{Y}_{t,P}^s,\ b_{i,t}^s \in \mathcal{B}_{i,t,P}^s \ \forall s \in \mathcal{S}, t \in \mathcal{T}, i \in \mathcal{I}$$

$$Z_t^{s,s'}, Z_t^{s',s} \in \mathcal{Z}_{t,P}^{s',s} \qquad \forall (s, s') \in \mathcal{L}_T^1, t \in \mathcal{T}$$

The set $\mathcal{N}_P^e$ consists of all tuples (s, s', t) for which equality non-anticipativity constraints link variables $x_t^{(\cdot)}$, $y_{t+1}^{(\cdot)}$ and $b_{i,t+1}^{(\cdot)}$ for scenarios s, s' in problem (P). Similarly, $\mathcal{N}_P^d$ represents the set of tuples (s, s', t) for which non-anticipativity constraints are applied in the form of disjunctions in problem (P). For example, in model $(P2_{ref})$,

$$\mathcal{N}_{P2_{ref}}^e = \left\{ (s, s', t) | (s, s') \in \mathcal{L}^0, t \in \mathcal{T}, t \le \mathbf{t}(s, s') \right\} \cup \left\{ (s, s', 0) | s, s' \in \mathcal{S}, s < s' \right\},$$

$$\mathcal{N}_{P2_{ref}}^d = \left\{ (s, s', t) | (s, s') \in \mathcal{L}_T^1, t \in \mathcal{T} \right\}.$$

$\mathcal{X}_{t,P}^s$, $\mathcal{Y}_{t,P}^s$, $\mathcal{B}_{i,t,P}^s$ and $\mathcal{Z}_{t,P}^{s,s'}$ represent the bounds and domain restrictions on variables x_t^s, y_t^s, $b_{i,t}^s$ and $Z_t^{s,s'}$ respectively, in model (P). The (relaxed) Lagrangean relaxation problem corresponding to (P) is obtained by relaxing all disjunctions for $(s, s', t) \in \mathcal{N}_P^d$ and replacing the symmetry conditions on variables $Z_t^{s,s'}$ with respect to (s, s') and the equality constraints corresponding to $(s, s', t) \in \mathcal{N}_P^e$ by penalty terms in the objective function. The corresponding relaxed Lagrangean dual is solved to obtain the lower bound for problem (P).

Based on the definitions of models (P) and $(P2_{RLR})$, an outline of the proposed algorithm is presented below. In the algorithm, $\mathcal{P}$ denotes the list of current problems together with the associated lower bounds, ϕ_{RLD}, while ϕ^{UB} represents the objective value of the best feasible solution obtained. For simplicity, we have assumed that all integer components of variables x_t^s and y_t^s correspond to binary variables.

Step 1 *Initialization:* $\phi^{UB} = \infty$, $\mathcal{P} = \left\{ P2_{ref} \right\}$.
Step 2 *Termination:* If $\mathcal{P} = \emptyset$, stop. Current best solution is optimal.

Step 3 *Node selection:* Select and delete P from $\mathcal{P}$. Solve relaxed Lagrangean dual of P to obtain the solution $(\hat{b}, \hat{w}, \hat{x}, \hat{y}, \hat{Z})$ with objective value $\phi_{RLD}(P)$. If the Lagrangean dual is infeasible, go to step 2. Set $\phi_{LB} = \phi_{RLD}(P)$.

Step 4 *Bounding:* If $\phi_{LB} \geq \phi^{UB}$, go to step 2 (This step can be carried out as soon as the value of the Lagrangean dual goes above ϕ^{UB}).
Apply a heuristic on solution $(\hat{b}, \hat{w}, \hat{x}, \hat{y}, \hat{Z})$ to generate feasible solution $(\bar{b}, \bar{w}, \bar{x}, \bar{y}, \bar{Z})$ with objective value $\bar{\phi}$. If heuristic is successful, let $\phi^{UB} := \min\left(\phi^{UB}, \bar{\phi}\right)$. Delete from $\mathcal{P}$ all problems P' with $\phi_{RLD}(P') \geq \phi^{UB}$.
If $\phi_{LB} \geq \phi^{UB}$, go to step 2.
else, go to step 5.

Step 5 *Branching:* Execute step (a) or (b).

(a) *On dualized equality constraints:* Select $(s_a, s_a', t_a) \in \mathcal{N}_P^e$. Create problems P_1, P_2 identical to P. Execute branching sub-step.

(b) *On relaxed disjunctions:* Select $(s_b, s_b', t_b) \in \mathcal{N}_P^d$. Create problems P_1, P_2, P_3 identical to P. Add restrictions $Z_{t_b}^{s_b,s_b'} = Z_{t_b}^{s_b',s_b} = True$ to P_1 and P_2 and $Z_{t_b}^{s_b,s_b'} = Z_{t_b}^{s_b',s_b} = False$ to P_3, respectively. For P_1 and P_2 update $\mathcal{N}_{(\cdot)}^e := \mathcal{N}_{(\cdot)}^e \cup (s_b, s_b', t_b)$, $\mathcal{N}_{(\cdot)}^d := \mathcal{N}_{(\cdot)}^d \setminus (s_b, s_b', t_b)$. Set $\mathcal{N}_{P_3}^d := \mathcal{N}_{P_3}^d \setminus (s_b, s_b', t_b)$. Add P_3 to $\mathcal{P}$. Select $(s_a, s_a', t_a) \in \mathcal{N}_{P_1}^e$. Execute branching sub-step.

Branching sub-step:
If $t_a = 0$, execute (ii) or (iii).
else if $t_a = T$, execute (i).
else, execute any one of (i), (ii) and (iii).

(i) Select component x_{l,t_a} of x_{t_a}.
If $x_{l,t_a} \in \{0, 1\}$ Add bounds $x_{l,t_a}^{s_a}, x_{l,t_a}^{s_a'} \geq 1$ to P_1 and $x_{l,t_a}^{s_a}, x_{l,t_a}^{s_a'} \leq 0$ to P_2.
else Add bounds $x_{l,t_a}^{s_a}, x_{l,t_a}^{s_a'} \geq \tilde{x}_{l,t_a}^{s_a,s_a'}$ to P_1 and $x_{l,t_a}^{s_a}, x_{l,t_a}^{s_a'} \leq \tilde{x}_{l,t_a}^{s_a,s_a'}$ to P_2.

(ii) Select component y_{l,t_a+1} of y_{t_a+1}.
If $y_{l,t_a+1} \in \{0, 1\}$ Add bounds $y_{l,t_a+1}^{s_a}, y_{l,t_a+1}^{s_a'} \geq 1$ to P_1 and $y_{l,t_a+1}^{s_a}, y_{l,t_a+1}^{s_a'} \leq 0$ to P_2.
else Add bounds $y_{l,t_a+1}^{s_a}, y_{l,t_a+1}^{s_a'} \geq \tilde{y}_{l,t_a+1}^{s_a,s_a'}$ to P_1 and $y_{l,t_a+1}^{s_a}, y_{l,t_a+1}^{s_a'} \leq \tilde{y}_{l,t_a+1}^{s_a,s_a'}$ to P_2.

(iii) Select variable b_{i,t_a+1}.
Add bounds $b_{i,t_a+1}^{s_a}, b_{i,t_a+1}^{s_a'} \geq 1$ to P_1 and $b_{i,t_a+1}^{s_a}, b_{i,t_a+1}^{s_a'} \leq 0$ to P_2.

Add P_1, P_2 to $\mathcal{P}$. Go to step 2.

Note that the solution of the Lagrangean dual in step 3 of the algorithm may not satisfy the relaxed disjunctions and the dualized equality constraints. In that case, feasible solutions can be generated in step 4 by applying problem-specific heuristics to the solution of the Lagrangean dual.

The branching step (step 5) partitions the feasible space by branching on violated equality constraints and disjunctions. The strategy used in step 5a for branching on dualized equality constraints is similar to that used by Caroe and Schultz [4]. Branching on equality constraints linking variables $b_{i,t}^{(\cdot)}$ (or the binary components of variables $x_t^{(\cdot)}$ or $y_t^{(\cdot)}$) across scenarios s, s' is based on the standard dichotomy branching strategy. When branching on constraint $x_{l,t}^s = x_{l,t}^{s'}$, where $x_{l,t}^{(\cdot)}$ is a continuous component of $x_t^{(\cdot)}$, the feasible space is partitioned about $\tilde{x}_{l,t}^{s,s'} = (p^s \hat{x}_{l,t}^s + p^{s'} \hat{x}_{l,t}^{s'})/(p^s + p^{s'})$, which is the mean value of variables $x_{l,t}^s$ and $x_{l,t}^{s'}$ in the solution of the Lagrangean dual. The same strategy is used for branching on equality constraints on continuous components of variables $y_t^{(\cdot)}$.

When branching on a relaxed disjunction corresponding to $(s, s', t) \in \mathcal{N}_P^d$ (step 5b), the feasible region is bifurcated into regions where $Z_t^{s,s'} = Z_t^{s',s} = True$ and $Z_t^{s,s'} = Z_t^{s',s} = False$, respectively. The set of dualized equality constraints on the up-branch ($Z_t^{s,s'} = Z_t^{s',s} = True$) is augmented by the set of equality constraints inside the disjunction corresponding to (s, s', t). Note that the solution of the relaxed Lagrangean dual of problem (P) may be such that $\hat{Z}_t^{s,s'} = \hat{Z}_t^{s',s} = True$. Thus, introducing the restriction $Z_t^{s,s'} = Z_t^{s',s} = True$ may not alter the solution of the Lagrangean dual. Thus, the first branch is further bifurcated to eliminate infeasibility in one of the violated equality constraints (see branching sub-step).

The problem at hand will govern the order in which the dualized equality constraints and the relaxed disjunctions are chosen for branching. Although we do not mention this step in the algorithm, logic inferencing (Hooker [10]) on Boolean and discrete variables can significantly impact the quality of the lower bounds. In this algorithm, constraints (4a), (5b), (17) and (18) can be used for logic inferencing on variables b and Z. For example, consider problem (P) with bound $b_{i_1,1}^{s_1} \leq 0$. We can use logic constraint (18) to infer that $Z_1^{s_1,s_2} = True$ for s_2 such that $\mathcal{D}(s_1, s_2) = \{i_1\}$. The set of equality constraints (to be dualized) can therefore be augmented as $\mathcal{N}_P^e := \mathcal{N}_P^e \cup (s_1, s_2, 1)$. Also, we can use equality constraint (4a) to infer that $b_{i_1,1}^s \leq 0 \; \forall s \in \mathcal{S}$. These changes to problem (P) may impact the value of the Lagrangean dual significantly. Therefore, it is important to use the logic inferencing step before the Lagrangean dual is solved in step 3 of the algorithm.

It should be noted that if some components of variables x_t and y_t are continuous, then some stopping criterion is needed to avoid infinite branching on these components. As explained by Caroe and Schultz [4], if the feasible region is bounded and if we branch parallel to the coordinate axes, then we can stop after the l_∞-diameter of the feasible sets of the sub-problems has fallen below a certain threshold. The algorithm is then guaranteed to converge finitely.

8. Results

In this section, we present qualitative and quantitative results for two specific instances of the problem described in section 3. All problems are solved using ILOG CPLEX 9.0 on a Pentium-IV, 2.4 GHz Linux machine.

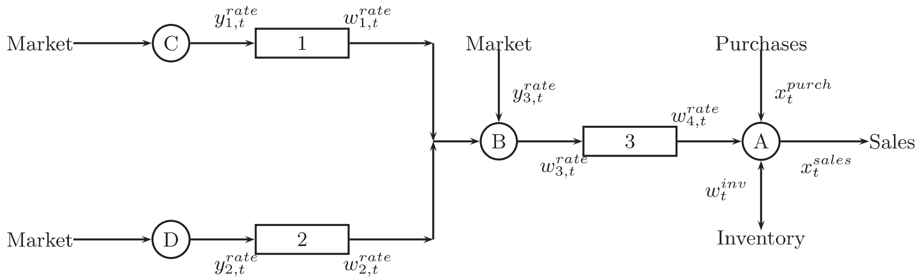


Fig. 2. Schematic representation of process network problem

8.1. Capacity expansion of process networks

The problem of capacity expansion in process networks is a specific example of the class of problems described in section 3. Figure 2 shows a process network that can be used to produce chemical A. Currently, A is being produced in unit 3 from chemical B, which is being purchased from the market. However, new technology is now available in the form of units 1 and 2 which can produce B from raw materials C and D, respectively. Chemicals C and D can be purchased from the market. If needed, A can also be purchased from the market. Also, inventory of chemical A can be maintained.

The demand for chemical A has to be satisfied in each time period t over a given time horizon $\mathcal{T}$. Decisions to be made in time period t include determining whether specific units should be operated in time period t or not (variables $b_{i,t} \in \{0, 1\}$ for $i = 1, 2, 3$), whether specific units should be installed or expanded (variables $y_{i,t}^{exp} \in \{0, 1\}$ for $i = 1, 2, 3$), the expansions in capacities of the units (variables $y_{i,t}^{QE}$ for $i = 1, 2, 3$), the feed flow rates (variables $y_{j,t}^{rate}$ for $j = 1, 2, 3$), and the purchase and sales to satisfy demand of A (variables x_t^{purch} and x_t^{sales}).

The yields (tons of product per ton of raw material) of units 1 and 2, represented by $\theta_{(\cdot)}$, are uncertain. Also, the future demands for chemical A, represented by ξ_t for $t \in \mathcal{T}$, are uncertain. The uncertainty in yield of a unit will be resolved only after that unit has been installed and operated for one time period. The uncertainty in demand in a specific time period is resolved automatically in that time-period.

The sequence of events in each time period is as follows. Decisions regarding which units to install or expand, the expansions in capacities of these units, which units to operate, and the feed flow rates are decided at the beginning of the time period. The resulting network configuration is then operated at the decided flow rates for the duration of the time period. The uncertainties in the demands for that specific time period and in the yields of units 1 and 2, if any of these units have been operated in this time period for the first time, are then resolved. Based on the observed realizations for the demand and yields, the sales and purchase decisions for that time period are made at the end of the time period.

$(CAPEXP)$ is the stochastic programming model corresponding to $(P2)$ for the capacity expansion problem described above. Please refer to the nomenclature section at the end of the paper for notation.

$(CAPEXP)$

$$\phi = \min \sum_{s \in \mathcal{S}} p^s w^{cost,s} \tag{30}$$

$$\text{s.t.} \quad w_{3,t}^{rate,s} = w_{1,t}^{rate,s} + w_{2,t}^{rate,s} + y_{3,t}^{rate,s} \qquad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{31}$$

$$w_t^{inv,s} = (w_{4,t}^{rate,s} + x_t^{purch,s} - x_t^{sales,s})\delta_t \qquad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{32}$$

$$w_{1,t}^{rate,s} = \theta_1^s \cdot y_{1,t}^{rate,s} \qquad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{33}$$

$$w_{2,t}^{rate,s} = \theta_2^s \cdot y_{2,t}^{rate,s} \qquad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{34}$$

$$w_{4,t}^{rate,s} = \theta_3^s \cdot w_{3,t}^{rate,s} \qquad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{35}$$

$$x_t^{sales,s} \geq \xi_t^s \qquad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{36}$$

$$w_{1,t}^{rate,s} \leq w_{1,t}^{Q,s} \qquad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{37}$$

$$w_{2,t}^{rate,s} \leq w_{2,t}^{Q,s} \qquad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{38}$$

$$w_{4,t}^{rate,s} \leq w_{3,t}^{Q,s} \qquad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{39}$$

$$w_{i,t}^{Q,s} = w_{i,t-1}^{Q,s} + y_{i,t}^{QE,s} \qquad \forall s \in \mathcal{S}, t \in \mathcal{T}, i = 1, 2, 3 \tag{40}$$

$$L_i^{QE} y_{i,t}^{exp,s} \leq y_{i,t}^{QE,s} \leq U_i^{QE} y_{i,t}^{exp,s} \qquad \forall s \in \mathcal{S}, t \in \mathcal{T}, i = 1, 2, 3 \tag{41}$$

$$L_1^{inflow} b_{1,t}^s \leq w_{1,t}^{rate,s} \leq U_1^{inflow} b_{1,t}^s \qquad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{42}$$

$$L_2^{inflow} b_{2,t}^s \leq w_{2,t}^{rate,s} \leq U_2^{inflow} b_{2,t}^s \qquad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{43}$$

$$L_3^{inflow} b_{3,t}^s \leq w_{4,t}^{rate,s} \leq U_3^{inflow} b_{3,t}^s \qquad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{44}$$

$$b_{i,t}^s \leq \sum_{\tau=1}^{t} y_{i,\tau}^{exp,s} \qquad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{45}$$

$$y_{i,t}^{exp,s} \leq b_{i,t}^s \qquad \forall s \in \mathcal{S}, t \in \mathcal{T}, i = 1, 2, 3 \tag{46}$$

$$\begin{aligned} w^{cost,s} = \quad & \sum_{t \in \mathcal{T}} \sum_{i=1}^{3} \left(FE_{i,t} y_{i,t}^{exp,s} + VE_{i,t} y_{i,t}^{QE,s} \right) + \sum_{t \in \mathcal{T}} \sum_{i=1}^{3} FO_{i,t} b_{i,t}^s \\ & + \sum_{t \in \mathcal{T}} \sum_{j=1}^{3} VO_{j,t}^y y_{j,t}^{rate,s} + \sum_{t \in \mathcal{T}} \sum_{j=1}^{4} VO_{j,t}^w w_{j,t}^{rate,s} \\ & + \sum_{t \in \mathcal{T}} \left(\alpha_t x_t^{purch,s} - \beta_t x_t^{sales,s} + \gamma_t w_t^{inv,s} \right) \quad \forall s \in \mathcal{S} \end{aligned} \tag{47}$$

$$b_{i,1}^s = b_{i,1}^{s'} \qquad \forall s, s' \in \mathcal{S}, s < s', i = 1, 2, 3 \tag{48a}$$

$$y_{i,1}^{exp,s} = y_{i,1}^{exp,s'} \qquad \forall s, s' \in \mathcal{S}, s < s', i = 1, 2, 3 \tag{48b}$$

$$y_{i,1}^{QE,s} = y_{i,1}^{QE,s'} \qquad \forall s, s' \in \mathcal{S}, s < s', i = 1, 2, 3 \tag{48c}$$

$$y_{j,1}^{rate,s} = y_{j,1}^{rate,s'} \qquad \forall s, s' \in \mathcal{S}, s < s', j = 1, 2, 3 \tag{48d}$$

$$x_t^{purch,s} = x_t^{purch,s'} \quad \forall (s,s') \in \mathcal{L}^0, t \in \mathcal{T}, t \le \mathbf{t}(s,s') \tag{49a}$$

$$x_t^{sales,s} = x_t^{sales,s'} \quad \forall (s,s') \in \mathcal{L}^0, t \in \mathcal{T}, t \le \mathbf{t}(s,s') \tag{49b}$$

$$b_{i,t+1}^{s} = b_{i,t+1}^{s'} \quad \forall (s,s') \in \mathcal{L}^0, t \in \mathcal{T}, t \le \mathbf{t}(s,s'), i = 1,2,3 \tag{49c}$$

$$y_{i,t+1}^{exp,s} = y_{i,t+1}^{exp,s'} \quad \forall (s,s') \in \mathcal{L}^0, t \in \mathcal{T}, t \le \mathbf{t}(s,s'), i = 1,2,3 \tag{49d}$$

$$y_{i,t+1}^{QE,s} = y_{i,t+1}^{QE,s'} \quad \forall (s,s') \in \mathcal{L}^0, t \in \mathcal{T}, t \le \mathbf{t}(s,s'), i = 1,2,3 \tag{49e}$$

$$y_{j,t+1}^{rate,s} = y_{j,t+1}^{rate,s'} \quad \forall (s,s') \in \mathcal{L}^0, t \in \mathcal{T}, t \le \mathbf{t}(s,s'), j = 1,2,3 \tag{49f}$$

$$\begin{bmatrix} & Z_t^{s,s'} & & \\ x_t^{purch,s} & = & x_t^{purch,s'} & \\ x_t^{sales,s} & = & x_t^{sales,s'} & \\ b_{i,t+1}^{s} & = & b_{i,t+1}^{s'} & \text{if } t \le T-1; i = 1,2,3 \\ y_{i,t+1}^{exp,s} & = & y_{i,t+1}^{exp,s'} & \text{if } t \le T-1; i = 1,2,3 \\ y_{i,t+1}^{QE,s} & = & y_{i,t+1}^{QE,s'} & \text{if } t \le T-1; i = 1,2,3 \\ y_{j,t+1}^{rate,s} & = & y_{j,t+1}^{rate,s'} & \text{if } t \le T-1; j = 1,2,3 \end{bmatrix} \vee \left[\neg Z_t^{s,s'}\right]$$

$$\forall (s,s') \in \mathcal{L}_T^1, t \in \mathcal{T} \tag{50}$$

$$Z_t^{s,s'} \Leftrightarrow \left[\bigwedge_{\tau=1}^{t} \left(\neg b_{i,\tau}^{s}\right)\right] \quad \forall (s,s') \in \mathcal{L}_T^1, t \in \mathcal{T}, \{i\} = \mathcal{D}(s,s') \tag{51}$$

$$\begin{array}{rlll} 0 \le & y_{j,t}^{rate,s} & \le U_j^{rate,y} & \forall s \in \mathcal{S}, t \in \mathcal{T}, j = 1,2,3 \\ 0 \le & w_{i,t}^{rate,s} & \le U_j^{rate,w} & \forall s \in \mathcal{S}, t \in \mathcal{T}, j = 1,2,3,4 \\ 0 \le & y_{i,t}^{QE,s} & \le U_i^{QE} & \forall s \in \mathcal{S}, t \in \mathcal{T}, i = 1,2,3 \\ 0 \le & w_{i,t}^{Q,s} & \le U_i^{Q} & \forall s \in \mathcal{S}, t \in \mathcal{T}, i = 1,2,3 \\ 0 \le & x_t^{purch,s} & \le U_t^{purch} & \forall s \in \mathcal{S}, t \in \mathcal{T} \\ 0 \le & x_t^{sales,s} & \le U_t^{sales} & \forall s \in \mathcal{S}, t \in \mathcal{T} \\ 0 \le & w_t^{inv,s} & \le U_t^{inv} & \forall s \in \mathcal{S}, t \in \mathcal{T} \\ & y_{i,t}^{exp,s} & \in \{0,1\} & \forall s \in \mathcal{S}, t \in \mathcal{T}, i = 1,2,3 \\ & b_{i,t}^{s} & \in \{0,1\} & \forall s \in \mathcal{S}, t \in \mathcal{T}, i = 1,2,3 \\ & Z_t^{s,s'} & \in \{True, False\} & \forall (s,s') \in \mathcal{L}_T^1, t \in \mathcal{T} \end{array}$$

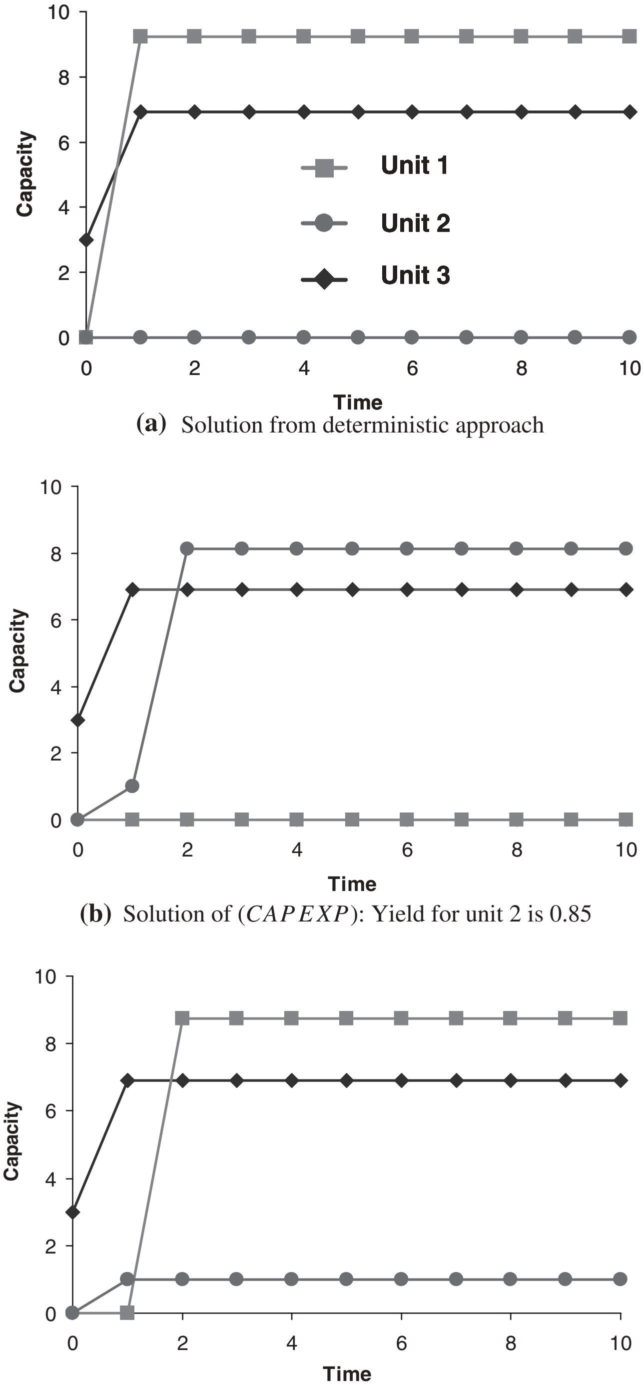
Expression (30) represents the objective of minimizing the overall expected cost. Equalities (31)-(32) correspond to mass balance constraints, while constraints (33)-(35) represent the yield relationships for the three process units. Constraint (36) forces the sales in time period t of scenario s to be greater than the demand in that time period and scenario. Capacity restrictions on the outflow from the three process units are applied by inequality constraints (37)-(39). The capacities of the three process units in time period t are computed by equality constraints (40). Inequalities (41) represent variable upper bounds on the capacity expansion decisions, while inequalities (42)-(44) represent variable upper bounds on the flow rates into process units. Constraint (45) states that a process unit can be operated in time period t only if it has been installed in or before time period t. Constraint (46) states that a process unit has to be operated in time period t if its capacity is expanded in that time period. The total cost in scenario s is computed in equation (47). Constraints (31)-(47) are applied for every scenario and correspond to constraints (3) in model $(P2)$. Decisions for different scenarios are linked together by non-anticipativity constraints (48)-(50). Constraints (48), (49), (50) and (51) correspond to constraints (4), (5), (17) and (18), respectively, in $(P2)$.

Example 1. Capacity expansion and operation decisions for the problem presented in Figure 2 have to be optimized over a time horizon of 10 time periods. Unit 3 is already operational with an existing capacity of 3 tons/hour and a known yield of 0.70. Possible realizations for the yield of unit 1 are 0.69 and 0.81, both possibilities being equally likely. Similarly, 0.65 and 0.85 are equally probable realizations for the yield of unit 2. Note that the mean yield for both these units is 0.75. However, the yield for unit 2 has greater variance. For the sake of simplicity, we assume that the demands are known deterministically, and hence, there is no exogenous uncertainty.

We compare the solution of the stochastic program $(CAPEXP)$ with that obtained from the deterministic approach. In the deterministic approach, the expected value problem is solved and the solution implemented partially till some uncertainty is resolved. The deterministic model is then updated and re-solved to obtain optimal decisions for the future. The solution of the expected value problem proposes installation of unit 1 and expansion of unit 3 in time period 1 (Fig. 3(a)). The approach proposes no more investments irrespective of the realization for the yield of unit 1. The expected cost for this solution is US $415,810. The solution of $(CAPEXP)$ proposes the expansion of unit 3 and the installation of unit 2 with a small capacity in time period 1 (Figures 3(b), 3(c)). Proposed investments in time period 2 are based on the realization for the yield of unit 2. Specifically, unit 2 should be expanded further in time period 2 if the yield for that unit is found to be 0.85 (Figure 3(b)). Otherwise, unit 1 should be installed in time period 2 (Figure 3(c)). The expected cost for this solution is US $403,718. Therefore, the value of stochastic solution for this problem is US $12,092.

This value arises because the proposed stochastic program is able to realize the "value of information" associated with the yield of unit 2 (the yield for unit 2 could be very high or very low). The solution proposes small investment to resolve the uncertainty in unit 2 and proposes appropriate recourse based on the realization for the yield of this unit. The reduction in expected cost in excess of 2.9% is a result of this flexibility provided by the stochastic program.

Table 1 compares the sizes of the big-M reformulations of the stochastic programs for the process network problem before and after Theorems 1 and 2 are applied. As can be seen, these theorems lead to a 32.49% reduction in the number of constraints. This in turn leads to 44.44% reduction in the CPU time required by ILOG CPLEX 9.0 to solve the problem.



(a) Solution from deterministic approach

(b) Solution of $(CAPEXP)$: Yield for unit 2 is 0.85

(c) Solution of $(CAPEXP)$: Yield for unit 2 is 0.65

Fig. 3. Solutions for Example 1

Table 1. MILP for Example 1: Effect of Theorems 1 and 2

	Before applying theorems	After applying theorems	% Reduction
Binary variables	240	240	0.00
Continuous variables	985	965	2.03
Constraints	3,853	2,601	32.49
CPU seconds	18	10	44.44

8.2. *Sizes problem*

The sizes problem (Jorjani et al. [13], Jonsbraten et al. [12]) is another specific example of the class of problems under consideration. In this problem, a production line has to meet the demand for a product in a set of different sizes, $\mathcal{I} = \{1, 2, \ldots, I\}$, in each time period $t \in \mathcal{T}$. If the demand for a particular size cannot be met in a specific time period, the deficit can be filled by the delivery of a bigger size. However, this involves a substitution cost. Other costs include fixed production costs for set-up of equipment for each size produced in each time period, variable inventory costs and variable production costs for each unit produced.

The demands, represented by ξ_t for time period $t \in \mathcal{T}$, are uncertain. The variable production costs, represented by θ_i for size $i \in \mathcal{I}$, remain constant over the time horizon but are also uncertain. The demand in time period t will be observed automatically in that time period. On the other hand, the uncertainty in variable production cost for size i, θ_i, will be resolved only when that size is produced for the first time. Thus, demand uncertainty is exogenous while the uncertainty in variable production costs is endogenous.

Decisions to be made in each time period include whether to produce size i or not (binary variables $b_{i,t}$), number of units of size i to be produced (variables $y_{i,t}$) and number of units of size i to be used to satisfy the demand for size i' (variables $x_{i,i',t}$). Production decisions $(b_{i,t}, y_t)$ are implemented at the beginning of time period t. This is followed by resolution of uncertainty in the demands for time period t and in the variable production costs for sizes that are produced for the first time in time period t. Finally, substitution decisions $(x_{i,i',t})$ to satisfy the demands in time period t are implemented at the end of that time period.

The stochastic programming model corresponding to $(P2)$ for the sizes problem is given below. Please refer to the nomenclature section at the end of the paper for notation.

$(SIZES)$

$$\phi = \min \sum_{s \in \mathcal{S}} p^s \sum_{t \in \mathcal{T}} \left(\sum_{i \in \mathcal{I}} \mu w_{i,t}^s + \sum_{i,i' \in \mathcal{I}, i \geq i'} \rho x_{i,i',t}^s + \sum_{i \in \mathcal{I}} \theta_i^s y_{i,t}^s + \sum_{i \in \mathcal{I}} \sigma b_{i,t}^s \right)$$

s.t.

$$y_{i,t}^s \geq {}^y L_i \cdot b_{i,t}^s \qquad \forall s \in \mathcal{S}, t \in \mathcal{T}, i \in \mathcal{I} \tag{52}$$

$$y_{i,t}^s \leq {}^y U_{i,t}^s \cdot b_{i,t}^s \qquad \forall s \in \mathcal{S}, t \in \mathcal{T}, i \in \mathcal{I} \tag{53}$$

$$\sum_{i \in \mathcal{I}} y_{i,t}^s \leq \alpha \qquad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{54}$$

$$\sum_{i \in \mathcal{I}, i \geq i'} x^s_{i,i',t} \geq \xi^s_{i',t} \qquad \forall s \in \mathcal{S}, t \in \mathcal{T}, i' \in \mathcal{I} \tag{55}$$

$$\sum_{\substack{\tau \in \mathcal{T}, \\ \tau \leq t}} \left(y^s_{i,\tau} - \sum_{i' \in \mathcal{I}, i \geq i'} x^s_{i,i',\tau} \right) = w^s_{i,t} \qquad \forall s \in \mathcal{S}, t \in \mathcal{T}, i \in \mathcal{I} \tag{56}$$

$$\sum_{i \in \mathcal{I}} w^s_{i,t} \leq \beta \qquad \forall s \in \mathcal{S}, t \in \mathcal{T} \tag{57}$$

$$b^s_{i,1} = b^{s'}_{i,1} \quad \forall s, s' \in \mathcal{S}, s < s', i \in \mathcal{I} \tag{58a}$$

$$y^s_{i,1} = y^{s'}_{i,1} \quad \forall s, s' \in \mathcal{S}, s < s', i \in \mathcal{I} \tag{58b}$$

$$x^s_{i,i',t} = x^{s'}_{i,i',t} \quad \forall (s, s') \in \mathcal{L}^0, t \in \mathcal{T}, t \leq \mathbf{t}(s, s'); i, i' \in \mathcal{I}, i \geq i' \tag{59a}$$

$$b^s_{i,t+1} = b^{s'}_{i,t+1} \quad \forall (s, s') \in \mathcal{L}^0, t \in \mathcal{T}, t \leq \mathbf{t}(s, s'); i \in \mathcal{I} \tag{59b}$$

$$y^s_{i,t+1} = y^{s'}_{i,t+1} \quad \forall (s, s') \in \mathcal{L}^0, t \in \mathcal{T}, t \leq \mathbf{t}(s, s'); i \in \mathcal{I} \tag{59c}$$

$$\begin{bmatrix} & Z^{s,s'}_t & \\ x^s_{i,i',t} & = & x^{s'}_{i,i',t} \ \forall i, i' \in \mathcal{I}, i \geq i' \\ b^s_{i,t+1} & = & b^{s'}_{i,t+1} \ \forall i \in \mathcal{I}; \text{if } t \leq T-1 \\ y^s_{i,t+1} & = & y^{s'}_{i,t+1} \ \forall i \in \mathcal{I}; \text{if } t \leq T-1 \end{bmatrix} \vee \left[\neg Z^{s,s'}_t \right] \forall (s, s') \in \mathcal{L}^1_T, t \in \mathcal{T} \tag{60}$$

$$Z^{s,s'}_t \Leftrightarrow \left[\bigwedge_{\tau=1}^{t} \left(\neg b^s_{i,\tau} \right) \right] \qquad \forall (s, s') \in \mathcal{L}^1_T, t \in \mathcal{T}, \{i\} = \mathcal{D}(s, s') \tag{61}$$

$$\begin{array}{rll} 0 \leq w^s_{i,t} & & \forall s \in \mathcal{S}, t \in \mathcal{T}, i \in \mathcal{I} \\ 0 \leq x^s_{i,i',t} & \leq {}^xU^s_{i,i',t} & \forall s \in \mathcal{S}, t \in \mathcal{T}; i, i' \in \mathcal{I}, i \geq i' \\ y^s_{i,t} & \in \mathcal{Y}^s_{i,t} & \forall s \in \mathcal{S}, t \in \mathcal{T}, i \in \mathcal{I} \\ b^s_{i,t} & \in \{0, 1\} & \forall s \in \mathcal{S}, t \in \mathcal{T}, i \in \mathcal{I} \\ Z^{s,s'}_t & \in \{True, False\} & \forall (s, s') \in \mathcal{L}^1_T, t \in \mathcal{T} \end{array}$$

Constraints (52) and (53) represent variable lower and upper bounds, respectively, on $y^s_{i,t}$. Constraint (54) enforces capacity restrictions on the total production in any time period. Constraint (55) ensures that demands for all sizes are satisfied in all time periods. Note that we assume that set $\mathcal{I}$ is ordered such that the size corresponding to index i is larger than the size corresponding to index i' if $i > i'$. Hence, delivery of size i can satisfy demand for size i'. Equation (56) computes the inventory of a size at the end of a time period, while constraint (57) enforces capacity restrictions on the the total inventory at the end of any time period. Constraints (58), (59), (60) and (61) are non-anticipativity constraints corresponding to constraints (4), (5), (17) and (18), respectively.

Note that the LP relaxation of the above model can be tightened by adding constraints (62)–(64), which correspond to dis-aggregation of variables $x^s_{i,i',t}$ into variables $w2^s_{i,i',\tau,t}$ for $\tau \leq t$ (see Krarup and Bilde [14]). Variable $w2^s_{i,i',\tau,t}$ represents the number of items of size i that are produced in time period τ but are used to satisfy demand of size i' ($\leq i$) in time period t ($\geq \tau$). Constraint (62) relates variables $x^s_{i,i',t}$ to dis-aggregated variables $w2^s_{i,i',\tau,t}$, while inequality (63) represents variable upper bounds on the dis-aggregated variables. Constraint (64) represents the condition that consumption of any size cannot exceed the production.

$$x^s_{i,i',t} = \sum_{\tau \in \mathcal{T}, \tau \leq t} w2^s_{i,i',\tau,t} \quad \forall s \in \mathcal{S}, t \in \mathcal{T}; i, i' \in \mathcal{I}, i \geq i' \tag{62}$$

$$w2^s_{i,i',\tau,t} \leq {}^{w2}U^s_{i,i',\tau,t} \cdot b^s_{i,\tau} \quad \forall s \in \mathcal{S}; \tau, t \in \mathcal{T}, \tau \leq t; \; i, i' \in \mathcal{I}, i \geq i' \tag{63}$$

$$\sum_{\substack{t \in \mathcal{T}, \\ t \geq \tau}} \sum_{\substack{i' \in \mathcal{I}, \\ i \geq i'}} w2^s_{i,i',\tau,t} \leq y^s_{i,\tau} \quad \forall s \in \mathcal{S}, \tau \in \mathcal{T}, i \in \mathcal{I} \tag{64}$$

We have developed an object-oriented implementation in C++ to solve the $(SIZES)$ model using the proposed branch and bound algorithm. In our implementation, the Lagrangean dual problem at each node is solved using a sub-gradient procedure (see Fisher [5]). This procedure is stopped if the lower bound does not improve for a pre-specified number of iterations, or if a total iteration limit is reached. At the root node, the Lagrange multipliers are initialized to zero. At all other nodes, the Lagrange multipliers are initialized to the optimal values at the parent node. Violated equality constraints are given higher priority for branching, compared to violated conditional non-anticipativity constraints (constraints (60) and (61)). The dualized equality constraint with the highest penalty is chosen for branching. If the corresponding penalty is less than a pre-specified value, the disjunction with the maximum violation is selected for branching.

A two-step heuristic procedure is used to generate feasible solutions at each node in the branch and bound tree. In the first step of this procedure, values for variables $b^s_{i,t}$ are fixed so that all non-anticipativity constraints involving these variables are satisfied. The MILP reformulation of model $(SIZES)$ is then solved in full space to obtain a feasible solution.

Let $\hat{b}^s_{i,t}$ represent the value of variable $b^s_{i,t}$ in the solution of the Lagrangean dual. The heuristic procedure fixes decisions $b^s_{i,t}$ in order of increasing t. Let $\bar{b}^s_{i,t}$ represent the value to which variable $b^s_{i,t}$ is fixed by the heuristic procedure. In order to fix decisions $b_{i,t}$, we identify the set of scenarios $s_1, s_2, \ldots, s_k$ such that constraint $b^{s_1}_{i,t} = b^{s_2}_{i,t} = \cdots = b^{s_k}_{i,t}$ has to be satisfied. For $j = 1, 2, \ldots, k$, we fix $\bar{b}^{s_j}_{i,t} = 1$ if $\sum_{j=1}^{k} \hat{b}^{s_j}_{i,t} - \sum_{j=1}^{k} \left(1 - \hat{b}^{s_j}_{i,t}\right) > 0$. Otherwise, $\bar{b}^{s_j}_{i,t}$ is set to 0 for $j = 1, 2, \ldots, k$.

If $t = 1$, constraint (58a) directly implies that the set $s_1, s_2, \ldots, s_k$ consists of all the scenarios in set $\mathcal{S}$. When $t > 1$, we first evaluate $\bar{Z}^{s,s'}_t$ for all $(s, s') \in \mathcal{L}^1_T$ by using logic constraint (61) together with the values of $\bar{b}^s_{(\cdot),t-1}$. The set $s_1, s_2, \ldots, s_k$ is then identified by combining equalities (59b) with the equality constraints that are enforced by disjunctions (60) when $Z^{s,s'}_t$ is fixed to the value of $\bar{Z}^{s,s'}_t$.

Example 2. For four test problems, we compare the performance of the proposed algorithm with the standard LP based branch and bound algorithm applied to the big-M reformulations of $(SIZES)$. The model specifications for these problems are presented

Table 2. Model specifications for sample problems in Example 2

Problem						
Name	I	T	S	Binary variables	Continuous variables	Constraints
A	5	5	16	400	4,977	13,927
B	4	5	27	540	6,103	22,160
C	6	6	32	1,152	18,081	52,581
D	7	7	64	3,136	60,993	181,939

Table 3. Computational results for Example 2

	Proposed Branch and Bound				LP based Branch and Bound					
	After t CPU seconds				After $2 \cdot t$ CPU seconds			After $10 \cdot t$ CPU seconds		
Problem Name	t	Nodes	Best sol. found	% gap	Nodes	Best sol. found	% gap	Nodes	Best sol. found	% gap
A	502	9	120,026	0.010	3,061	120,044	0.111	26,490	120,026	0.025
B	430	4	112,608	0.010	708	112,623	0.258	6,899	112,621	0.157
C	11,546	15	144,054	0.065	10,963	144,169	0.411	79,844	144,154	0.332
D	13,507	3	245,930	0.038	15,591	246,029	0.247	71,151	246,026	0.237

Table 4. Comparison of lower bounds at root node in Example 2

Problem	Proposed algorithm		ILOG CPLEX	
Name	Lower bound at root node	% gap	Lower bound at root node	% gap
A	120,003	0.019	119,770	0.213
B	112,595	0.012	112,240	0.327
C	143,742	0.217	143,466	0.408
D	245,403	0.214	245,341	0.239

in Table 2, where I, T and S represent the number of sizes, time periods and scenarios, respectively. Note that we only consider endogenous uncertainty in problems A-D. Also, because the magnitudes of the demands in each time period are fairly large ($\approx 10,000$), variables $y^s_{i,t}$ (number of units produced) and $x^s_{i,i',t}$ (number of units substituted) are treated as continuous variables.

Table 3 compares the status of the proposed algorithm after t CPU seconds with that of the standard LP based branch and bound algorithm (ILOG CPLEX 9.0) after $2 \cdot t$ CPU seconds and $10 \cdot t$ CPU seconds. Clearly, the LP based branch and bound is not able to obtain the same optimality gap even after one order of magnitude more CPU time than that taken by the proposed algorithm. Also, in comparable CPU time the proposed branch and bound algorithm generates better feasible solutions. This can be attributed in part to the tighter lower bounds obtained from the relaxed Lagrangean dual. Table 4 compares the lower bounds generated by the two algorithms at the root node of the branch and bound tree. The gaps are calculated relative to the best solution found (column 4 of Table 3).

Finally, it should be noted that while the results presented in Table 3 are very encouraging, more extensive computational experiments will be conducted in the future to verify the effectiveness of the proposed algorithm.

9. Conclusions

In this paper, we have addressed a class of stochastic programs where the optimization decisions determine the times when the uncertainties in various parameters will be resolved. We have extended the stochastic programming modeling framework by incorporating the interaction between the optimization decisions and the information discovery process in the form of conditional non-anticipativity constraints which are applied inside disjunctions. We have presented theoretical properties that lead to significant reduction in the size of the proposed model. Most importantly, we have shown that the non-anticipativity constraints need to be applied for scenarios s and s' only if the scenarios either differ exclusively in realizations for exogenous uncertain parameters, or differ exclusively in the realization of *one* endogenous uncertain parameter. We have also presented a Lagrangean duality based branch and bound algorithm to solve the model.

This approach has been applied to a capacity expansion problem in chemical networks and to the manufacturing related sizes problem. Results show that the inclusion of the option of getting information at the cost of investments leads to significant improvements in the quality of the solution. Computational results presented for the proposed branch and bound algorithm show that more than one order of magnitude reduction in solution time can be achieved over the standard LP based branch and bound algorithm.

A. Proof of Proposition 1

Suppose vectors $\hat{b}^s, \hat{b}^{s'}, \hat{Z}^{s,s'}$ satisfy (9)–(11).

Proof (Proof of (a)). Consider $t \in \mathcal{T}, t \leq \hat{t}$. Let

$$\bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}^s_{i,\tau} \right) \right] = True.$$

$$\Rightarrow \bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t'} \left(\neg \hat{b}^s_{i,\tau} \right) \right] = True \quad \forall t' \in \mathcal{T}, t' \leq t$$

Since $t \leq \hat{t}$ and $(\hat{b}^s, \hat{b}^{s'}, \hat{Z}^{s,s'})$ satisfies constraint (11), therefore,

$$\hat{Z}^{s,s'}_\tau = True \quad \forall \tau \in \mathcal{T}, \tau \leq t.$$

Using constraints (9) and (10) we get

$$\hat{b}^s_{i,\tau} = \hat{b}^{s'}_{i,\tau} \qquad \forall i \in \mathcal{I}, \tau \in \mathcal{T}, \tau \leq t+1$$

Thus,

$$\bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}^s_{i,\tau} \right) \right] \Rightarrow \left[\hat{b}^s_{i,\tau} = \hat{b}^{s'}_{i,\tau} \quad \forall i \in \mathcal{I}, \tau \in \mathcal{T}, \tau \leq t+1 \right].$$

□

Proof (Proof of (b)). Suppose

$$\bigwedge_{i\in\mathcal{D}(s,s')}\left[\bigwedge_{\tau=1}^{\hat{t}+1}\left(\neg\hat{b}^{s}_{i,\tau}\right)\right]=True.$$

$$\Rightarrow\bigwedge_{i\in\mathcal{D}(s,s')}\left[\bigwedge_{\tau=1}^{\hat{t}}\left(\neg\hat{b}^{s}_{i,\tau}\right)\right]=True.$$

Now using the result of part (a), we get,

$$\hat{b}^{s}_{i,\tau}=\hat{b}^{s'}_{i,\tau}\quad\forall i\in\mathcal{I},\tau\in\mathcal{T},\tau\le\hat{t}+1$$

Thus, for $t=\hat{t}+1$,

$$\bigwedge_{i\in\mathcal{D}(s,s')}\left[\bigwedge_{\tau=1}^{t}\left(\neg\hat{b}^{s}_{i,\tau}\right)\right]\Rightarrow\left[\hat{b}^{s}_{i,\tau}=\hat{b}^{s'}_{i,\tau}\quad\forall i\in\mathcal{I},\tau\in\mathcal{T},\tau\le t\right].$$

□

Proof (Proof of (c)). Consider $t\in\mathcal{T}, t\le\hat{t}+1$.
(⇒):
Suppose,

$$\bigwedge_{i\in\mathcal{D}(s,s')}\left[\bigwedge_{\tau=1}^{t}\left(\neg\hat{b}^{s}_{i,\tau}\right)\right]=True$$

From statements (a) and (b) of this proposition,

$$\hat{b}^{s}_{i,\tau}=\hat{b}^{s'}_{i,\tau}\quad\forall i\in\mathcal{I},\tau\in\mathcal{T},\tau\le t$$

Therefore,

$$\bigwedge_{i\in\mathcal{D}(s,s')}\left[\bigwedge_{\tau=1}^{t}\left(\neg\hat{b}^{s'}_{i,\tau}\right)\right]=True$$

Thus,

$$\bigwedge_{i\in\mathcal{D}(s,s')}\left[\bigwedge_{\tau=1}^{t}\left(\neg\hat{b}^{s}_{i,\tau}\right)\right]\Rightarrow\bigwedge_{i\in\mathcal{D}(s,s')}\left[\bigwedge_{\tau=1}^{t}\left(\neg\hat{b}^{s'}_{i,\tau}\right)\right]\quad\forall t\in\mathcal{T},t\le\hat{t}+1\tag{A.1}$$

(⇐):
We prove the contra-positive of the converse. Suppose

$$\bigwedge_{i\in\mathcal{D}(s,s')}\left[\bigwedge_{\tau=1}^{t}\left(\neg\hat{b}^{s}_{i,\tau}\right)\right]=False$$

Thus, there exists (i, τ) such that $i \in \mathcal{D}(s, s')$, $\tau \in \mathcal{T}$, $\tau \leq t$ for which $\hat{b}^s_{i,\tau} = 1$. Define

$$\tau^* = \min_{\tau} \left\{ \tau | \tau \in \mathcal{T}, \tau \leq t \text{ such that } \exists i \in \mathcal{D}(s, s'), \hat{b}^s_{i,\tau} = 1 \right\}$$

and

$$i^* \in \left\{ i | i \in \mathcal{D}(s, s'), \hat{b}^s_{i,\tau^*} = 1 \right\}$$

Proof (Case 1: $\tau^* = 1$). By definition of i^* and τ^*, $\hat{b}^s_{i^*,\tau^*} = 1$. Thus, using equality constraint (9) we get

$$\hat{b}^s_{i^*,\tau^*} = \hat{b}^{s'}_{i^*,\tau^*} = 1$$

Since $\tau^* = 1 \leq t$ and $i^* \in \mathcal{D}(s, s')$, therefore

$$\bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}^{s'}_{i,\tau} \right) \right] = False \tag{A.2}$$

$\square$

Proof (Case 2: $\tau^* > 1$:). By definition of τ^*, $\hat{b}^s_{i,\tau} = 0 \quad \forall i \in \mathcal{D}(s, s'), \tau \in \mathcal{T}, \tau \leq \tau^* - 1$. Thus,

$$\bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{\tau^*-1} \left(\neg \hat{b}^{s}_{i,\tau} \right) \right] = True \tag{A.3}$$

Now, since $1 < \tau^* \leq t \leq \hat{t} + 1$ and $\tau^* \in \mathcal{T} = \{1, 2, \ldots, T\}$, therefore $\tau^* - 1 \in \mathcal{T}$ and $1 \leq \tau^* - 1 \leq \hat{t}$. Then, using the result of part (a) together with (A.3), we get

$$\hat{b}^s_{i,\tau} = \hat{b}^{s'}_{i,\tau} \quad \forall i \in \mathcal{I}, \tau \in \mathcal{T}, \tau \leq \tau^* \tag{A.4}$$

From definition of (i^*, τ^*), $\hat{b}^s_{i^*,\tau^*} = 1$. Using (A.4) we get $\hat{b}^{s'}_{i^*,\tau^*} = 1$. Now, since $\tau^* \leq t$ and $i^* \in \mathcal{D}(s, s')$, therefore,

$$\bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}^{s'}_{i,\tau} \right) \right] = False \tag{A.5}$$

$\square$

Combining (A.2) and (A.5), we get

$$\neg \left(\bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}^{s}_{i,\tau} \right) \right] \right) \Rightarrow \neg \left(\bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}^{s'}_{i,\tau} \right) \right] \right) \quad \forall t \in \mathcal{T}, t \leq \hat{t} + 1 \tag{A.6}$$

Thus, from (A.1) and (A.6) we get

$$\bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}^{s}_{i,\tau} \right) \right] \Leftrightarrow \bigwedge_{i \in \mathcal{D}(s,s')} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}^{s'}_{i,\tau} \right) \right] \quad \forall t \in \mathcal{T}, t \leq \hat{t} + 1$$

$\square$

Proof (Proofs for (d) and (e)). Follow directly by combining (c) with (a) and (b) respectively. □

B. Lemma

Lemma 1. *Consider scenarios $s_0, s_1, s_2, \ldots, s_r \in \mathcal{S}$ and time period $\hat{t} \in \mathcal{T}$, $\hat{t} \leq T-1$. Suppose vectors $\hat{b}$, $\hat{Z}$ are such that for $k \in \{0, 1, \ldots, r-1\}$,*

(i) If $s_k < s_{k+1}$, then sub-vectors $\hat{b}^{s_k}, \hat{b}^{s_{k+1}}, \hat{Z}^{s_k, s_{k+1}}$ satisfy constraints (9)–(11) for $(s, s') = (s_k, s_{k+1})$

(ii) If $s_{k+1} < s_k$, then sub-vectors $\hat{b}^{s_{k+1}}, \hat{b}^{s_k}, \hat{Z}^{s_{k+1}, s_k}$ satisfy constraints (9)–(11) for $(s, s') = (s_{k+1}, s_k)$.

Then, vector $\hat{b}$ satisfies

$$\bigwedge_{k=0}^{r-1}\left[\bigwedge_{\tau=1}^{t}\left(\bigwedge_{i\in\mathcal{D}(s_k,s_{k+1})}\left(\neg\hat{b}_{i,\tau}^{s_k}\right)\right)\right] \Leftrightarrow \bigwedge_{k=0}^{r-1}\left[\bigwedge_{\tau=1}^{t}\left(\bigwedge_{i\in\mathcal{D}(s_k,s_{k+1})}\left(\neg\hat{b}_{i,\tau}^{s_0}\right)\right)\right] \quad \forall t \in \mathcal{T}, t \leq \hat{t}+1$$

Proof. Consider $t \in \mathcal{T}$ such that $t \leq \hat{t}+1$. We will first show that

$$\bigwedge_{\tau=1}^{t}\left(\bigwedge_{i\in\mathcal{D}(s_k,s_{k+1})}\left(\neg\hat{b}_{i,\tau}^{s_k}\right)\right) \Rightarrow \left[\hat{b}_{i,\tau}^{s_k} = \hat{b}_{i,\tau}^{s_{k+1}} \quad \forall i \in \mathcal{I}, \tau \in \mathcal{T}, \tau \leq t\right] \quad k = 0, 1, \ldots, r-1 \qquad \text{(B.1)}$$

(B.1) will be used repeatedly in the proof of the Lemma.

If $k \in \{0, 1, \ldots, r-1\}$ and $s_k < s_{k+1}$, we can use condition (i) to infer that sub-vectors $\hat{b}^{s_k}, \hat{b}^{s_{k+1}}, \hat{Z}^{s_k, s_{k+1}}$ satisfy (9)–(11) for $(s, s') = (s_1, s_2)$. Hence, we can combine statements (a) and (b) of Proposition 1 to get (B.1). On the other hand, if $s_{k+1} < s_k$, we can use condition (ii) to infer that sub-vectors $\hat{b}^{s_{k+1}}, \hat{b}^{s_k}, \hat{Z}^{s_{k+1}, s_k}$ satisfy constraints (9)–(11) for $(s, s') = (s_{k+1}, s_k)$. Hence, we can combine statements (d) and (e) of Proposition 1 to infer (B.1). Thus, (B.1) holds for $k \in \{0, 1, \ldots, r-1\}$ irrespective of whether $s_k < s_{k+1}$ or $s_{k+1} < s_k$. Now we prove the Lemma under consideration.

(⇒): Suppose

$$\bigwedge_{k=0}^{r-1}\left[\bigwedge_{\tau=1}^{t}\left(\bigwedge_{i\in\mathcal{D}(s_k,s_{k+1})}\left(\neg\hat{b}_{i,\tau}^{s_k}\right)\right)\right] = True$$
$$\Rightarrow \bigwedge_{\tau=1}^{t}\left(\bigwedge_{i\in\mathcal{D}(s_k,s_{k+1})}\left(\neg\hat{b}_{i,\tau}^{s_k}\right)\right) = True \ \forall k \in \{0, 1, \ldots, r-1\}$$

Using (B.1), we get

$$\hat{b}_{i,\tau}^{s_k} = \hat{b}_{i,\tau}^{s_{k+1}} \quad \forall i \in \mathcal{I}, \tau \in \mathcal{T}, \tau \leq t, k \in \{0, 1, \ldots, r-1\}$$

Thus,

$$\hat{b}^{s_0}_{i,\tau} = \hat{b}^{s_1}_{i,\tau} = \hat{b}^{s_2}_{i,\tau} = \cdots = \hat{b}^{s_{r-1}}_{i,\tau} = \hat{b}^{s_r}_{i,\tau} \quad \forall i \in \mathcal{I}, \tau \in \mathcal{T}, \tau \leq t$$

Using this in the hypothesis, we get

$$\bigwedge_{k=0}^{r-1}\left[\bigwedge_{\tau=1}^{t}\left(\bigwedge_{i\in\mathcal{D}(s_k,s_{k+1})}\left(\neg\hat{b}^{s_0}_{i,\tau}\right)\right)\right] = True$$

(⇐): Suppose

$$\bigwedge_{k=0}^{r-1}\left[\bigwedge_{\tau=1}^{t}\left(\bigwedge_{i\in\mathcal{D}(s_k,s_{k+1})}\left(\neg\hat{b}^{s_0}_{i,\tau}\right)\right)\right] = True$$

$$\Rightarrow \bigwedge_{\tau=1}^{t}\left(\bigwedge_{i\in\mathcal{D}(s_k,s_{k+1})}\left(\neg\hat{b}^{s_0}_{i,\tau}\right)\right) = True \quad \forall k \in \{0, 1, \ldots, r-1\} \tag{B.2}$$

Therefore, $\bigwedge_{\tau=1}^{t}\left(\bigwedge_{i\in\mathcal{D}(s_0,s_1)}\left(\neg\hat{b}^{s_0}_{i,\tau}\right)\right) = True$. Hence, from (B.1) we get

$$\hat{b}^{s_0}_{i,\tau} = \hat{b}^{s_1}_{i,\tau} \quad \forall i \in \mathcal{I}, \tau \in \mathcal{T}, \tau \leq t \tag{B.3}$$

Replacing $\hat{b}^{s_0}_{i,\tau}$ by $\hat{b}^{s_1}_{i,\tau}$ in (B.2) we can infer that $\bigwedge_{\tau=1}^{t}\left(\bigwedge_{i\in\mathcal{D}(s_1,s_2)}\left(\neg\hat{b}^{s_1}_{i,\tau}\right)\right) = True$. Once again, using (B.1) we can infer that

$$\hat{b}^{s_1}_{i,\tau} = \hat{b}^{s_2}_{i,\tau} \quad \forall i \in \mathcal{I}, \tau \in \mathcal{T}, \tau \leq t \tag{B.4}$$

Hence, combining (B.2), (B.3) and (B.4) we can infer that $\bigwedge_{\tau=1}^{t}\left(\bigwedge_{i\in\mathcal{D}(s_2,s_3)}\left(\neg\hat{b}^{s_2}_{i,\tau}\right)\right) = True$. We can continue the argument to show that,

$$\hat{b}^{s_0}_{i,\tau} = \hat{b}^{s_1}_{i,\tau} = \hat{b}^{s_2}_{i,\tau} = \cdots = \hat{b}^{s_r}_{i,\tau} \quad \forall i \in \mathcal{I}, \tau \in \mathcal{T}, \tau \leq t$$

Using this in the hypothesis we get

$$\bigwedge_{k=0}^{r-1}\left[\bigwedge_{\tau=1}^{t}\left(\bigwedge_{i\in\mathcal{D}(s_k,s_{k+1})}\left(\neg\hat{b}^{s_k}_{i,\tau}\right)\right)\right] = True$$

□

C. Proof of Proposition 2

Proof. Suppose vectors $\hat{b}, \hat{x}, \hat{y}, \hat{Z}$ satisfy (i)–(iii) in the condition of the proposition. Consider scenarios s_a, s_b such that

$$\begin{aligned}\theta_i^{s_a} &= \theta_i^{s_b} \; i \in \{1, 2, \ldots, I-r\}\\ \theta_i^{s_a} &\neq \theta_i^{s_b} \; i \in \{I-r+1, I-r+2, \ldots, I\}\end{aligned}$$

Thus, $\mathcal{D}(s_a, s_b) = \{I-r+1, I-r+2, \ldots, I\}$ and $|\mathcal{D}(s_a, s_b)| = r$. We consider the case where $1 \leq r = |\mathcal{D}(s_a, s_b)| \leq I$. Without loss of generality, we assume that $s_a < s_b$ and that s_a, s_b differ in realizations for the *last* r endogenous parameters.

Hence, $(s_a, s_b) \in \mathcal{L}^{1+}$. We will prove that vectors $\hat{b}, \hat{x}, \hat{y}, \hat{Z}$ satisfy constraints (15)–(16) for (s_a, s_b, t) such that $t \in \mathcal{T}, t \leq \mathbf{t}(s_a, s_b)$. Since $(s_a, s_b) \in \mathcal{L}^{1+}$, this will establish that $(\hat{b}, \hat{x}, \hat{y}, \hat{Z})$ satisfies (15)–(16) for all (s, s', t) such that $(s, s') \in \mathcal{L}^{1+}, t \in \mathcal{T}, t \leq \mathbf{t}(s, s')$. The result will follow.

Case 1. $\mathbf{t}(s_a, s_b) < T$

Scenarios s_a, s_b differ in realizations of r endogenous parameters. We construct r "intermediate" scenarios indexed by variable indices $s_1, s_2, \ldots, s_r$. These scenarios are derived from scenario s_a by progressively changing the realization of one of these r endogenous parameters to the corresponding realization in scenario s_b. For example, scenario s_1 is identical to scenario s_a except that $\theta_{I-r+1}^{s_1} = \theta_{I-r+1}^{s_b} \neq \theta_{I-r+1}^{s_a}$. Similarly, scenario s_2 is identical to scenario s_a except that $\theta_{I-r+1}^{s_2} = \theta_{I-r+1}^{s_b} \neq \theta_{I-r+1}^{s_a}$ and $\theta_{I-r+2}^{s_2} = \theta_{I-r+2}^{s_b} \neq \theta_{I-r+2}^{s_a}$. Scenarios $s_1, s_2, \ldots, s_r$ are identical to scenario s_a in terms of realizations for all exogenous parameters. Mathematically,

$$\begin{array}{llll}\theta_i^{s_1} = \theta_i^{s_a} \;\forall i \in \mathcal{I} \setminus \{I-r+1\}, & \theta_i^{s_1} = \theta_i^{s_b} \;\forall i \in \{I-r+1\}\\ \theta_i^{s_2} = \theta_i^{s_a} \;\forall i \in \mathcal{I} \setminus \{I-r+1, I-r+2\}, & \theta_i^{s_2} = \theta_i^{s_b} \;\forall i \in \{I-r+1, I-r+2\}\\ \vdots \quad \vdots \quad \vdots & \vdots \quad \vdots \quad \vdots\\ \theta_i^{s_r} = \theta_i^{s_a} \;\forall i \in \mathcal{I} \setminus \{I-r+1, \ldots, I\}, & \theta_i^{s_r} = \theta_i^{s_b} \;\forall i \in \{I-r+1, \ldots, I\}\end{array}$$

and

$$\xi_t^{s_k} = \xi_t^{s_a} \quad \forall t \in \mathcal{T}, k \in \{1, 2, \ldots, r\}$$

Since we choose the set of scenarios as $\Xi \times (\times_{i \in \mathcal{I}} \Theta_i)$, the set of scenarios includes scenarios corresponding to all possible combinations of realizations for the endogenous parameters for any realization of the vector of exogenous parameters, $\xi = (\xi_1, \xi_2, \ldots, \xi_T)$. Thus, the constructed scenarios corresponding to the indices $s_1, s_2, \ldots, s_k$ belong to the set of scenarios. Therefore, we can choose indices $s_1, s_2, \ldots, s_k$ such that $s_1, s_2, \ldots, s_k \in \mathcal{S}$.

By construction,

$$\mathcal{D}(s_a, s_1) = \{I-r+1\}, \mathcal{D}(s_1, s_2) = \{I-r+2\}, \ldots, \mathcal{D}(s_{r-1}, s_r) = \{I\}$$

and,

$$\mathcal{D}(s_r, s_b) = \emptyset.$$

Therefore,

$$|\mathcal{D}(s_a, s_1)| = |\mathcal{D}(s_1, s_2)| = \cdots = |\mathcal{D}(s_{r-1}, s_r)| = 1$$

while

$$|\mathcal{D}(s_r, s_b)| = 0.$$

Also, by construction, $\xi_t^{s_a} = \xi_t^{s_k}\ \forall t \in \mathcal{T}, k \in \{1, 2, \ldots, r\}$. Hence,

$$\mathbf{t}(s_a, s_1) = \mathbf{t}(s_1, s_2) = \cdots = \mathbf{t}(s_{r-1}, s_r) = T \text{ and}$$

$$\mathbf{t}(s_r, s_b) = \mathbf{t}(s_a, s_b).$$

Consider scenarios s_a and s_1.

Case 1a. $s_a < s_1$

Since $|\mathcal{D}(s_a, s_1)| = 1$ and $\mathbf{t}(s_a, s_1) = T$, therefore $(s_a, s_1) \in \mathcal{L}_T^1$. Thus, from condition (iii) of the statement of this proposition, we can infer that vectors $\hat{b}, \hat{x}, \hat{y}, \hat{Z}$ satisfy (15)–(16) for (s_a, s_1, t), where $t \in \mathcal{T}$. Since $\mathcal{D}(s_a, s_1) = \{I - r + 1\}$, from (15) and (16) we have,

$$\left[\bigwedge_{\tau=1}^{t}\left(\neg \hat{b}_{I-r+1,\tau}^{s_a}\right)\right] \Rightarrow \begin{bmatrix} \hat{x}_t^{s_a} & = \hat{x}_t^{s_1} & \\ \hat{b}_{i,t+1}^{s_a} & = \hat{b}_{i,t+1}^{s_1}\ \forall i \in \mathcal{I} & \text{if } t \leq T-1 \\ \hat{y}_{t+1}^{s_a} & = \hat{y}_{t+1}^{s_1} & \text{if } t \leq T-1 \end{bmatrix} \quad \forall t \in \mathcal{T} \quad \text{(C.1)}$$

Case 1b. $s_1 < s_a$

In this case, $(s_1, s_a) \in \mathcal{L}_T^1$. Therefore from condition (iii) of the statement of this proposition, we can infer that

$$\left[\bigwedge_{\tau=1}^{t}\left(\neg \hat{b}_{I-r+1,\tau}^{s_1}\right)\right] \Rightarrow \begin{bmatrix} \hat{x}_t^{s_a} & = \hat{x}_t^{s_1} & \\ \hat{b}_{i,t+1}^{s_a} & = \hat{b}_{i,t+1}^{s_1}\ \forall i \in \mathcal{I} & \text{if } t \leq T-1 \\ \hat{y}_{t+1}^{s_a} & = \hat{y}_{t+1}^{s_1} & \text{if } t \leq T-1 \end{bmatrix} \quad \forall t \in \mathcal{T} \quad \text{(C.2)}$$

Also, since $s_1 < s_a$, vectors $\hat{b}, \hat{x}, \hat{y}, \hat{Z}$ satisfy equality constraint (13a) for $(s_1, s_a, t = 1)$. Thus, sub-vectors $\hat{b}^{s_1}, \hat{b}^{s_a}, \hat{Z}^{s_1,s_a}$ satisfy equality constraint (13a) for $(s, s', t) = (s_1, s_a, 1)$ and constraints (15)–(16) for $(s, s') = (s_1, s_a), t \in \mathcal{T}$. Thus, sub-vectors $\hat{b}^{s_1}, \hat{b}^{s_a}, \hat{Z}^{s_1,s_a}$ satisfy constraints (9)–(11) for $(s, s', \hat{t}) = (s_1, s_a, T-1)$. Thus, using $\mathcal{D}(s_a, s_1) = \{I - r + 1\}$ in statement (c) of Proposition 1, we get

$$\left[\bigwedge_{\tau=1}^{t}\left(\neg \hat{b}_{I-r+1,\tau}^{s_1}\right)\right] \Leftrightarrow \left[\bigwedge_{\tau=1}^{t}\left(\neg \hat{b}_{I-r+1,\tau}^{s_a}\right)\right] \quad \forall t \in \mathcal{T}, t \leq \hat{t} + 1 = T \quad \text{(C.3)}$$

Combining (C.1)–(C.3), we obtain that irrespective of whether $s_a < s_1$ or $s_1 < s_a$,

$$\left[\bigwedge_{\tau=1}^{t}\left(\neg \hat{b}_{I-r+1,\tau}^{s_a}\right)\right] \Rightarrow \begin{bmatrix} \hat{x}_t^{s_a} & = \hat{x}_t^{s_1} & \\ \hat{b}_{i,t+1}^{s_a} & = \hat{b}_{i,t+1}^{s_1}\ \forall i \in \mathcal{I} & \text{if } t \leq T-1 \\ \hat{y}_{t+1}^{s_a} & = \hat{y}_{t+1}^{s_1} & \text{if } t \leq T-1 \end{bmatrix} \quad \forall t \in \mathcal{T} \quad \text{(C.4)}$$

Since

$$|\mathcal{D}(s_1, s_2)| = |\mathcal{D}(s_2, s_3)| = \cdots = |\mathcal{D}(s_{r-1}, s_r)| = 1,$$

$$\mathbf{t}(s_1, s_2) = \mathbf{t}(s_2, s_3) = \cdots = \mathbf{t}(s_{r-1}, s_r) = T,$$

and

$$\mathcal{D}(s_1, s_2) = \{I - r + 2\}, \mathcal{D}(s_2, s_3) = \{I - r + 3\}, \ldots, \mathcal{D}(s_{r-1}, s_r) = \{I\},$$

we can use the same logic as above to prove that

$$\left[\bigwedge_{\tau=1}^{t}\left(\neg \hat{b}^{s_1}_{I-r+2,\tau}\right)\right] \Rightarrow \begin{bmatrix} \hat{x}^{s_1}_t & = \hat{x}^{s_2}_t & \\ \hat{b}^{s_1}_{i,t+1} & = \hat{b}^{s_2}_{i,t+1} & \forall i \in \mathcal{I} \text{ if } t \leq T-1 \\ \hat{y}^{s_1}_{t+1} & = \hat{y}^{s_2}_{t+1} & \text{if } t \leq T-1 \end{bmatrix} \quad \forall t \in \mathcal{T} \tag{C.5}$$

$$\vdots \quad \vdots \quad \vdots$$

$$\left[\bigwedge_{\tau=1}^{t}\left(\neg \hat{b}^{s_{r-1}}_{I,\tau}\right)\right] \Rightarrow \begin{bmatrix} \hat{x}^{s_{r-1}}_t & = \hat{x}^{s_r}_t & \\ \hat{b}^{s_{r-1}}_{i,t+1} & = \hat{b}^{s_r}_{i,t+1} & \forall i \in \mathcal{I} \text{ if } t \leq T-1 \\ \hat{y}^{s_{r-1}}_{t+1} & = \hat{y}^{s_r}_{t+1} & \text{if } t \leq T-1 \end{bmatrix} \quad \forall t \in \mathcal{T} \tag{C.6}$$

Also, $|\mathcal{D}(s_r, s_b)| = 0$. Therefore, depending on whether $s_r < s_b$ or $s_b < s_r$, either $(s_r, s_b) \in \mathcal{L}^0$ or $(s_b, s_r) \in \mathcal{L}^0$. In either case, from condition (ii) of this proposition we have

$$\left.\begin{array}{ll} \hat{x}^{s_r}_t & = \hat{x}^{s_b}_t \\ \hat{b}^{s_r}_{i,t+1} & = \hat{b}^{s_b}_{i,t+1} \ \forall i \in \mathcal{I} \\ \hat{y}^{s_r}_{t+1} & = \hat{y}^{s_b}_{t+1} \end{array}\right\} \quad \forall t \in \mathcal{T}, t \leq \mathbf{t}(s_r, s_b)$$

Now since $\mathbf{t}(s_r, s_b) = \mathbf{t}(s_a, s_b)$, therefore replacing $\mathbf{t}(s_r, s_b)$ by $\mathbf{t}(s_a, s_b)$ we get

$$\left.\begin{array}{ll} \hat{x}^{s_r}_t & = \hat{x}^{s_b}_t \\ \hat{b}^{s_r}_{i,t+1} & = \hat{b}^{s_b}_{i,t+1} \forall i \in \mathcal{I} \\ \hat{y}^{s_r}_{t+1} & = \hat{y}^{s_b}_{t+1} \end{array}\right\} \quad \forall t \in \mathcal{T}, t \leq \mathbf{t}(s_a, s_b) \tag{C.7}$$

Combining (C.4)–(C.7), we get

$$\bigwedge_{\tau=1}^{t}\left[\left(\neg \hat{b}^{s_a}_{I-r+1,\tau}\right) \wedge \left(\neg \hat{b}^{s_1}_{I-r+2,\tau}\right) \wedge \left(\neg \hat{b}^{s_2}_{I-r+3,\tau}\right) \wedge \cdots \wedge \left(\neg \hat{b}^{s_{r-1}}_{I,\tau}\right)\right] \Rightarrow$$
$$\begin{bmatrix} \hat{x}^{s_a}_t & = \hat{x}^{s_1}_t & = \cdots = \hat{x}^{s_r}_t & = \hat{x}^{s_b}_t & \\ \hat{b}^{s_a}_{i,t+1} & = \hat{b}^{s_1}_{i,t+1} & = \cdots = \hat{b}^{s_r}_{i,t+1} & = \hat{b}^{s_b}_{i,t+1} & \forall i \text{ if } t \leq T-1 \\ \hat{y}^{s_a}_{t+1} & = \hat{y}^{s_1}_{t+1} & = \cdots = \hat{y}^{s_r}_{t+1} & = \hat{y}^{s_b}_{t+1} & \text{if } t \leq T-1 \end{bmatrix} \forall t \in \mathcal{T}, t \leq \mathbf{t}(s_a, s_b) \tag{C.8}$$

We complete the proof using Lemma 1. To streamline the use of Lemma 1, we will refer to s_a as s_0. By construction, $|\mathcal{D}(s_k, s_{k+1})| = |\mathcal{D}(s_{k+1}, s_k)| = 1$ and $\mathbf{t}(s_k, s_{k+1}) = \mathbf{t}(s_{k+1}, s_k) = T$ for $k \in \{0, 1, \ldots, r-1\}$. Thus, from conditions (i) and (iii) of this proposition,

(a) If $s_k < s_{k+1}$, then $(s_k, s_{k+1}) \in \mathcal{L}_T^1$. Hence, vectors $\hat{b}, \hat{x}, \hat{y}, \hat{Z}$ satisfy equality constraint (13a) for $(s, s', t) = (s_k, s_{k+1}, 1)$ and constraints (15)–(16) for $(s, s') = (s_k, s_{k+1}), t \in \mathcal{T}$.
(b) If $s_{k+1} < s_k$, then $(s_{k+1}, s_k) \in \mathcal{L}_T^1$. Hence, vectors $\hat{b}, \hat{x}, \hat{y}, \hat{Z}$ satisfy equality constraint (13a) for $(s, s', t) = (s_{k+1}, s_k, 1)$ and constraints (15)–(16) for $(s, s') = (s_{k+1}, s_k), t \in \mathcal{T}$.

Thus, choosing $\hat{t} = T - 1$ for $k \in \{0, 1, \ldots, r-1\}$,

(a) If $s_k < s_{k+1}$, then sub-vectors $\hat{b}^{s_k}, \hat{b}^{s_{k+1}}, \hat{Z}^{s_k, s_{k+1}}$ satisfy constraints (9)–(11) for $(s, s') = (s_k, s_{k+1})$.
(b) If $s_{k+1} < s_k$, then sub-vectors $\hat{b}^{s_{k+1}}, \hat{b}^{s_k}, \hat{Z}^{s_{k+1}, s_k}$ satisfy constraints (9)–(11) for $(s, s') = (s_{k+1}, s_k)$.

Thus, using Lemma 1 for $\hat{t} = T - 1$ together with

$$\mathcal{D}(s_a, s_1) = \{I - r + 1\}, \mathcal{D}(s_1, s_2) = \{I - r + 2\},$$
$$\mathcal{D}(s_2, s_3) = \{I - r + 3\}, \ldots, \mathcal{D}(s_{r-1}, s_r) = \{I\},$$

and $s_0 \equiv s_a$, we get

$$\bigwedge_{\tau=1}^{t} \left[\left(\neg \hat{b}^{s_a}_{I-r+1,\tau}\right) \wedge \left(\neg \hat{b}^{s_1}_{I-r+2,\tau}\right) \wedge \left(\neg \hat{b}^{s_2}_{I-r+3,\tau}\right) \wedge \cdots \wedge \left(\neg \hat{b}^{s_{r-1}}_{I,\tau}\right)\right]$$
$$\Leftrightarrow \bigwedge_{\tau=1}^{t} \left[\left(\neg \hat{b}^{s_a}_{I-r+1,\tau}\right) \wedge \left(\neg \hat{b}^{s_a}_{I-r+2,\tau}\right) \wedge \left(\neg \hat{b}^{s_a}_{I-r+3,\tau}\right) \wedge \cdots \wedge \left(\neg \hat{b}^{s_a}_{I,\tau}\right)\right] \quad \forall t \in \mathcal{T} \tag{C.9}$$

Combining (C.8) and (C.9) together with the fact that $\mathcal{D}(s_a, s_b) = \{I - r + 1, I - r + 2, \ldots, I\}$, we get

$$\bigwedge_{i \in \mathcal{D}(s_a, s_b)} \left[\bigwedge_{\tau=1}^{t} \left(\neg \hat{b}^{s_a}_{i,\tau}\right)\right]$$
$$\Rightarrow \begin{bmatrix} \hat{x}^{s_a}_t &= \hat{x}^{s_b}_t & \\ \hat{b}^{s_a}_{i,t+1} &= \hat{b}^{s_b}_{i,t+1} \ \forall i \in \mathcal{I} & \text{if } t \le T-1 \\ \hat{y}^{s_a}_{t+1} &= \hat{y}^{s_b}_{t+1} & \text{if } t \le T-1 \end{bmatrix} \quad \forall t \in \mathcal{T}, t \le \mathbf{t}(s_a, s_b) \tag{C.10}$$

which is simply the re-statement of constraints (15)–(16) for scenario pair (s_a, s_b).

Case 2. $\mathbf{t}(s_a, s_b) = T$

The line of reasoning used above can be used in this case also. However, by construction above, $\mathcal{D}(s_r, s_b) = \emptyset$ and $\mathbf{t}(s_a, s_r) = T$. Since $\mathbf{t}(s_a, s_b) = T$ in this case, therefore we will have $\mathbf{t}(s_r, s_b) = T$ and $\mathcal{D}(s_r, s_b) = \emptyset$. Thus, scenarios s_r, s_b are identical; *i.e.*,

$$\left(\xi^{s_r}_1, \xi^{s_r}_2, \ldots, \xi^{s_r}_T, \theta^{s_r}_1, \theta^{s_r}_2, \ldots, \theta^{s_r}_I\right) \equiv \left(\xi^{s_b}_1, \xi^{s_b}_2, \ldots, \xi^{s_b}_T, \theta^{s_b}_1, \theta^{s_b}_2, \ldots, \theta^{s_b}_I\right).$$

Thus, we can replace s_r by s_b in the above proof and use (C.4), (C.5) and (C.6) to obtain (C.8) (without using (C.7)). As in the previous case, Lemma 1 can then be used to obtain (C.10).

Since all we assumed about s_a, s_b was that $|\mathcal{D}(s_a, s_b)| \geq 1$, we can say that vectors $\hat{b}, \hat{x}, \hat{y}, \hat{Z}$ satisfy constraints (15)–(16) for all (s, s', t) such that $(s, s') \in \mathcal{L}^{1+}, t \in \mathcal{T}, t \leq \mathbf{t}(s, s')$. □

Acknowledgments The authors are thankful to Amr El-Bakry and Eric Mulkay from ExxonMobil Upstream Research Company for providing guidance in the implementation of the branch and bound algorithm.

References

1. Ahmed, S.: Strategic planning under uncertainty: Stochastic integer programming approaches. Ph.D. thesis, University of Illinois at Urbana-Champaign, 2000
2. Balas, E.: Disjunctive programming and a hierarchy of relaxations for discrete optimization problems. SIAM Journal on Algebraic and Discrete Methods **6** (3), 466–486 (1985)
3. Birge, J.R.: Stochastic programming computation and applications. INFORMS J. Comput. **9**, 111–133 (1997)
4. Caroe, C.C., Schultz, R.: Dual decomposition in stochastic integer programming. Oper. Res. Lett. **24**, 37–45 (1999)
5. Fisher, M.L.: The lagrangian-relaxation method for solving integer programming-problems. Management Science **27** (1), 1–18 (1981)
6. Goel, V., Grossmann, I.E.: A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. Computers and Chemical Engineering **28** (8), 1409–1429 (2004)
7. Grossmann, I. E.: Review of nonlinear mixed-integer and disjunctive programming techniques. Optimization and Engineering **3**, 227–252 (2002)
8. Guignard, M., Kim, S.: Lagrangean decomposition - A model yielding stronger lagrangean bounds. Mathematical Programming **39** (2), 215–228 (1987)
9. Held, H., Woodruff, D.L.: Heuristics for multi-stage interdiction of stochastic networks. Journal of Heuristics (Submitted for publication)
10. Hooker, J.: Logic-based methods for optimization: combining optimization and constraint satisfaction. New York : John Wiley and Sons 2000
11. Jonsbraten, T.W.: Optimization models for petroleum field exploitation. Ph.D. thesis, Norwegian School of Economics and Business Administration (1998)
12. Jonsbraten, T.W., Wets, R.J.B., Woodruff, D.L.: A class of stochastic programs with decision dependent random elements. Annals of Operations Research **82**, 83–106 (1998)
13. Jorjani, S., Scott, C.H., Woodruff, D.L.: Selection of an optimal subset of sizes. International Journal of Production Research **37** (16), 3697–3710 (1999)
14. Krarup, J., Bilde, O.: Plant location, set covering and economic lot size: an O(mn)-algorithm for structured problems. In: Collatz L., Wetterling W., (eds.); Optimierung bei graphentheoretischen and ganzzahligen Problemen. Vol. 36 of International Series of Numerical Mathematics. Birkhauser Verlag, Basel, 155–180 (1977)
15. Nemhauser, G.L., Wolsey, L.A.: Integer and combinatorial optimization. Wiley, New York, English Series: Wiley-Interscience series in discrete mathematics and optimization 1988
16. Pflug, G.: On-line optimization of simulated markovian processes. Math. Oper. Res. **15** (3), 381–395 (1990)
17. Rockafellar, R., Wets, R. J.-B.: Scenarios and policy aggregation in optimization under uncertainty. Math. Oper. Res. **16**, 119–147 (1991)
18. Ruszczynski, A.: Decomposition methods in stochastic programming. Math. Programming (Ser. B) **79**, 333–353 (1997)
19. Sahinidis, N.V.: Optimization under uncertainty: State-of-the-art and opportunities. Comput. Chem. Eng. **28** (6–7), 971–983 (2004)
20. Schultz, R.: Stochastic programming with integer variables. Mathematical Programming **97** (1–2), 285–309 (2003)
21. Turkay, M., Grossmann, I.E.: Disjunctive programming techniques for the optimization of process systems with discontinuous investment costs multiple size regions. Ind. Eng. Chem. Res. **35**, 2611–2623 (1996)
22. Viswanath, K., Peeta, S., Salman, S.: Investing in the links of a stochastic network to minimize expected shortest path length (Working paper)