EUROPA-UNIVERSITÄT VIADRINA FRANKFURT (ODER)

Economic Effects of Mobile Technologies on Operations of Sales Agents

Grigori Pichtchoulov

Knut Richter

European University Viadrina Frankfurt (Oder)

Department of Business Administration and Economics

Discussion Paper No. 226

2005

ISSN 1860 0921

Economic Effects of Mobile Technologies on Operations of Sales Agents

Grigori Pichtchoulov and Knut Richter

European University Viadrina, Frankfurt (Oder), Germany

January 2005

Abstract

In the presented paper we introduce an approach to assess particular economic effects which may arise with bringing mobile technologies into the field of sales and distribution. The research problem posed here comprises quite a special case where sales operations of a company are carried by its sales representatives, which may count as a resource allocation problem. We apply stochastic programming methodology to model the agents' multistage decision making in a distribution system with uncertain customer demands, and exemplify a potential improvement in the company's overall performance when mobile facilities are utilized for making decisions. We provide finally an efficient computational algorithm that delivers optimal decision making with and without mobile technologies, and computes the expected overall performance in both cases, for any configuration of a distribution system. Some computational results are presented.

Contents

1 Introduction

Distribution of goods is realized in many industries by means of representatives who travel through assigned territories, promote the products and services to customers and execute sales operations, while the goods sold by the representatives are normally being delivered to the customers later on. With that the efficiency of representatives' operations determines to a great degree the efficiency of a company in distributing its goods. Due to many practical reasons, like, for example, a large number of customers and/or their geographical dispersion, the distributing company may have to assign different territories or customer groups to several different representatives, which then carry their work independently of each other. Then an efficient allocation of goods to the representatives who then have to efficiently distribute these goods between the customers may become an important issue under demand uncertainty, and we focus in this paper on potential improvements which mobile technologies might deliver there being introduced into the company's decision making. Namely, enabling a centralised and real-time coordination of the representatives' operations, the mobile facilities may enable a better distribution of goods between the company's customers.

For that we consider a company which delivers a particular product to a number of predetermined customers, and model its distribution operations in two different environments: first, without any presence of mobile facilities in the company's decision making, and then, in the second case, making use of them. To see what consequences will it have for attaining the company's goals becomes in fact our task.

1.1 Decision making with no mobile facilities

We can represent the operating environment of our company in the following way. The company delivers some good to its customers from a central warehouse on a regular basis — once a week. This good can be material as well as immaterial, i.e. also service. We make the following assumptions regarding the customer demand:

- the individual weekly demand of each customer is uncertain and represented by a random variable;
- the random variables of individual demands are independent and identically distributed;
- their probability distribution is known to us (and to the company);
- at the beginning of each week there is a constant stock s reserved by the company in advance to serve the customers during the week;
- the customer demand becomes realized first after the allocation of the weekly stock is done;

The delivery operations of the company have one very important issue: before a weekly physical delivery goes on tour, the demands must become known to the company. We assume further that each individual demand can be learnt only on the spot by visiting the customer, and that this job is carried out by the company' sales representatives. The company assigns different geographical regions to different representatives.

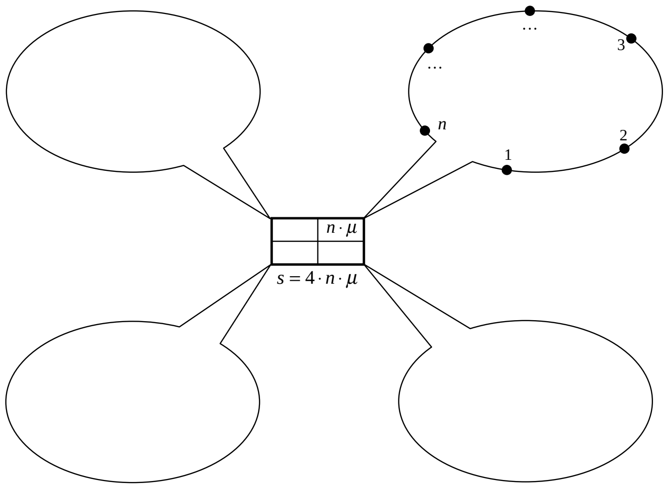


Figure 1: The company serves 4 districts with n customers in each.

We make the following assumptions concerning the work of the sales agents:

- each agent makes a round trip through his district at the beginning of the week and visits the assigned customers one by one;
- a customer's demand becomes known to the agent first by his arrival to the customer;
- having learnt this customer's actual demand, the agent makes a decision on the quantity to be delivered to this customer, and secures it by an appropriate contract with him;
- after that he drives to the next customer on the tour, and so on.

The following issue becomes important: even though the customer demands are not known in advance, still each agent should know right before the trip, what quantity of the good he has at his disposal. To secure this the company has to divide the total stock "virtually" between the agents, what it does according to some predefined rule, like, for example: the agents receive their quotas according to the expected total demands in their districts.

Since not every customer demand can be met by an agent fully, then we have to raise the following question: What should be a general rule for an agent while allocating the good for a given customer on the tour? Let us now formulate the assumptions regarding this decision making:

- the company has the objective: each customer's demand should be satisfied as fully as possible;
- while allocating the good for the given customer, an agent counts on the following issues:
    1. his available (not yet allocated) virtual stock,
    2. this customer's demand, and
    3. the total expected demand of his customers who hasn't been served yet;
- finally, the allocated (and secured) quantities are not subject to any later change.

So we may represent the company's objective in the following way: having in total N customers to serve, whose individual demands in a given week occur to be $\xi_1, \ldots, \xi_N$, and having met their demands in amounts $x_1, \ldots, x_N$, we express the quality of this week's entire decision $\mathbf{x}$ as

$$G(\mathbf{x}|\xi) \; = \; \left[\frac{x_1}{\xi_1} \; + \; \ldots \; + \; \frac{x_N}{\xi_N} \right] \cdot \frac{1}{N} \; , \tag{1}$$

where any $\xi_i = 0$ implies $x_i = 0$ and the corresponding fraction equals 1 by definition.

Let us make the following notes here:

- we have expressed by (1) the *overall customer satisfaction* in the given week, which we will also call this week's *overall performance*;
- the objective of the company's weekly decision making becomes the maximization of this overall performance.

The chosen measure of the overall customer satisfaction corresponds to the view in which the logistical customer service is defined as "the state in which customer needs, wants and expectations, through the transaction cycle, are met or exceeded, resulting in repurchase and continuing loyalty." [39, p. 390]. As a company tries to reach an optimal balancing between inventory costs and the customer service level, the chosen performance measure expresses the resulting degree of the product availability for a given time period with respect to each single customer, over all customers.

We would like to omit the constant factor $1/N$ in the company's objective since this is only a scaling factor. Since there is no interaction between the sales representatives assumed, the highest possible overall service level may be reached only through each agent's maximum performance. We will assume that all agents follow the same decision rules and so are equally good in decision making.

Referring to Figure 1 as an example and considering the 1st district with n customers, we see that the corresponding agent has to efficiently distribute his virtual stock $s_1 = n \cdot \mu$ with the objective of maximizing his *individual performance*:

$$\frac{x_1}{\xi_1} \; + \; \ldots \; + \; \frac{x_n}{\xi_n} \quad \longrightarrow \quad \max$$

We would like to make the following remarks:

- visiting the customers one by one, the agent performs a multi-stage decision making;
- the decisions on later stages are dependent on the earlier ones, since all stages share the same virtual stock of this agent;
- this implies that in some circumstances a relatively large demand of a particular customer should not be satisfied fully in order to provide an acceptable servicing of the customers in the next stages;
- from the other side, restricting the service for this customer too much deals with a risk that the future stages occur to have nevertheless a lower demand, and the virtual stock will be then not exhausted fully, what lowers the overall service level too;

- so an excessive as well a moderate policy on each decision stage can both be inefficient, and we raise the following general question:

Q1: How must the agent's optimal decision making look? Or, in other words, what should be his optimal strategy?

Let us now proceed to the presentation of the alternative environment which involves the use of mobile facilities intended to improve the overall quality of decision making.

1.2 Decision Making with Mobile Technologies

From now we let all representatives make use of mobile computer devices which are able to establish a connection with a company's stationary server at any time and from any customer location and upload the information on each allocated quantity to the server. If each representative will follow this policy then this will provide the company with the real-time data on how the stock is being allocated to the customers. This implies that the company will always have an up-to-date information on (a) how many customers are still awaiting an allocation, and (b) what fraction of the total stock has not been allocated by now.

Now let the mobile devices of the representatives be able not only to upload the data but also to retrieve this global up-to-date information to their mobile devices any time they need. With such capability the company doesn't need to divide the total stock between the representatives anymore, since they become able now to make allocations from the common stock.

Let us summarize the above in form of a new mobile policy for the representatives: upon arriving to a customer and learning his demand each representative retrieves from the central computer the information on the currently available stock and the number of the customers pending in the system, and tries to make such an allocation to the current customer that maximizes the overall system's performance. After making a decision he uploads it immediately on the central computer.

May these changes help to improve the overall decision making of the representatives? One can expect that they might, namely due to the following considerations:

- in the non-mobile environment any non-exhausted quota of one representative could not be shared by another one whose quota occurred to be not sufficient this week;
- in the mobile environment the representatives share the same (common) stock, what might lead to better allocations;
- they work in fact cooperatively in this mobile environment.

First of all, let us raise the following question:

Q2: How should the cooperative optimal decision making look?

And we come finally to the question

Q3: Which operating mode — with or without mobile facilities — is expected to deliver a better overall customer satisfaction? How such advantage, if any, does express?

We will answer these question through statement of mathematical models for both environments and analysing their optimal solutions.

1.3 Backlogging

The assumptions we made by now for both mobile and non-mobile decision making miss the following important point: what happens with uncovered customer demands? Do they get lost or backlogged to be covered at a later time? We suppose that letting them get lost would be rather an unrealistic assumption: in case when a representative has a particularly low stock at final stages of his trip, he has then almost nothing to offer to the final customers, but will still "curiously" figure out their demands.

And we extend our assumptions with the following one: if a representative doesn't meet any customer demand fully, then the missing merchandise must be delivered to this customer next week. Thus the company will order/produce for the next week additionally a certain (deterministically known) amount which will cover this week's total backlog.

Note that we make our model with this assumption to some extent closer to the concept of "available-to-promise" (see section 2.4). Namely, in case when a particular customer demand cannot be met fully according to the optimal allocation policy, it means that we are not able to promise to deliver the whole amount to the customer this week, but we can rather schedule the delivery of the required quantity in two portions: the 1st – this week, and the 2nd – next week. Also, if any customer demand occurs to be so extraordinary large that the company will not be able to supply the backlogged demand either in the coming week fully, then it can be scheduled to be delivered by parts week after week.

Lastly, we assume that the backlog doesn't incur any extra costs.

According to the declared above company's objective (1), the company is striving to provide each customer with a maximal possible allocation this week, what, in turn, means to minimize each customer's backlog. Thus, our assumption on backlogging doesn't require any modification of the company's objective, while is giving more practical sense to our model. The way the optimal decisions have to be made stays with that unchanged.

Thus, keeping in mind that backlogging takes place — and fitting our model to the ATP philosophy, we can nevertheless omit its explicit consideration in our forthcoming modeling.

Before we proceed to modeling, let us make a brief overview of the known research in related areas and also review the practical issues that might concern our problem field.

2 Related Research and Practical Issues

2.1 Vehicle Routing and Inventory-Routing Problems: Route Sales

Distribution of goods with deterministic demands was extensively studied by the research community over the last decades within the framework of Vehicle Routing Problems (VRPs). This classical extension of the Traveling Salesman Problem can be defined in its simplest form as follows: given a number of geographically dispersed customers with known demands, one has to construct a set of routes for identical vehicles of a limited capacity such that each customer is visited exactly once, all demands are fully met, vehicle capacities become never exceeded and the total travel distance minimized. For reviews of this research see for example [22, 26, 7, 4].

While the demand variability is not captured by the classical VRP, many real-life problems have to deal with it. The stochastic extensions of the VRP are studied in the research literature since the late 70's. Among them, the most extensively studied *stochastic vehicle routing problem* (SVRP) assumes the actual customer demands not to be known in advance but known to follow given probability distributions. With this information, one constructs the routes with the objective of minimizing the expected total distance travelled, taking into consideration potential *route failures* and *route breaks* for replenishing a vehicle at the depot. As a vehicle follows then its route and gradually reveals the actual demands (sooner or later - depending on assumptions), reoptimizations of the current route become possible (up to which extent - depending on the policy chosen), what the objective function has to account for, too. For more detailed presentations see [14, 6, 7, 19, 46, 44, 27].

As an example of real distribution problems where instances of the SVRP find their application, the following ones may be considered: delivery of petroleum products, loading the banks' automatic teller machines, or waste collection [37, p. 146], [46].

Yet another direction for extending the VRP has arisen due to the following considerations. An optimization of the transportation costs alone was not sufficient in many centralized distribution applications where retail outlets play the role of customers' locations and so the incurred inventory costs become an integral part of the system-wide costs. This need for an integrated coordination of inventory control and transportation planning in a single-warehouse-multiple-retailer system gave birth to a number of *inventory-routing problems* which extend the VRP in the following way: given a single depot with identical vehicles of a limited capacity and a number of retail outlets which face external demands at constant rates, one has to schedule optimal deliveries for replenishing the outlets and determine the corresponding vehicle routes so that the overall costs including those incurred by the retail inventory are minimized [2, 3, 9]. Inventory-routing problems with stochastic external demands were approached one of the first by Federgruen and Zipkin in [17] and Federgruen et al. in [16] where the objective functions incorporate the expectations of inventory and shortage costs. The recent papers by Kleywegt et al. [25] and Adelman [1] review the research done in this area.

We can make the following notes here. One may see that one of the basic assumptions made in the SVRP (as well as in the inventory-routing problems) is the *full delivery*: the customer's demand must be met fully if this can be done by the vehicle. This assumption excludes the cases when a limited resource has to be divided between the customers and so not every demand can be met fully.

Another important assumption of the SVRP and the inventory-routing problems is that they consider such distribution systems in which *Route Sales* occur to be the operating mode, what assumes that the vehicle carries the goods on board and distributes (sells) them as it follows its individual route. Thereby any re-allocation of the limited good between different vehicles (territories) becomes hard or costly to implement, even though it might appear worth doing as the customer demands become gradually revealed as the vehicles follow the routes.

2.2 Pre-Sell concept

In fact, we try to investigate in this paper such distribution operations which are free of the assumptions mentioned above and correspond to a greater extent to the *Pre-Sell* concept, in which the sales and the delivery operations are fully separated [2], what assumes that the sales are first arranged by the sales representatives with the customers before the deliveries will be executed by the company. Anily and Federgruen formulate this principle in the context of the inventory-routing applications as follows: "each salesman is required to visit the outlets in his region periodically in a given sequence determining replenishment quantities" [2]. We go a bit further in our assumptions and require that the quantities which a representative determines in such sequential manner cannot be changed or reconsidered later. So we are interested in the optimal determination of these quantities under presence of some remaining and not yet known demands having a limited total supply at our disposal.

We would like to mention here a note made by Yang et al. about a possible relaxation of the full delivery assumption in the SVRP context [46, p. 100]. Treating the vehicle's stock on board as a limited resource and wanting to serve the demand of remaining customers too, one might decide not to meet the demand of the current customer fully. The authors do not give this policy a further attention though.

Sequential decision making under uncertainty is discussed in the research literature dedicated to various applied areas, like, for example, target shooting [30], oil exploration [28], agriculture [29], portfolio management [38].

Our research can be reasoned by the current development of mobile technologies which, being introduced into the company's business processes, may support its employees in their decision making. The Pre-Sell distribution, reported as decreasing in the past decades [2, p. 93] due to a variety of advantages of Route Sales systems, is becoming now more popular due to availability of wireless devices which "allow field reps and management to do a better job of meeting demand" [21]. New software applications allowing the presell representatives to have an efficient access to vast amounts of corporate data via wireless handhelds are being at the time developed and applied. This all has "increased preselling activity and the resulting division between presales and delivery activities" [21]. The same source reports on estimations of about 25 to 30 percent of routes in the soft drink industry moving to presell. A much better and a real-time data exchange between a field representative and the company has a potential of improving the overall company's operating, since it may receive the customer orders from the representatives in real time (and make a better planning) and also provide all relevant information to the representative necessary for his actions.

Communication over phone and, even later, over mobile phones was not able to provide the same efficiency of data exchange between the company and a representative as which handheld computers now provide, since they give a faster and a "sharper" access to a broader corporate information [36].

Many other industries employ presell representatives as well, which, according to several occupational guides, typically have to travel through assigned territories, market, promote and sell products to manufacturers and wholesale and retail buyers, check existing stock, analyze, discuss and determine client's needs, inform about prices and availability, negotiate the sale, estimate delivery dates [15, 45].

A more recent online source [31] reports on a business project which two leading technology companies initiated to offer solutions for mobile workforce management in the utility industry, by means of which "...joint clients can optimize all types of field work, using the same asset management system that they use in their generation plants and facilities ... optimally schedule and wirelessly dispatch work to field crews, who use the solution to manage and record their work, and provide real-time status updates to and from the field."

We would like to treat this as a justification for considering the distribution operations under the assumptions discussed above and to investigate potential improvements which mobile technologies could provide in this area. As one may already see, planning of delivery operations does not come into our view, rather an optimal allocation of the goods to the customers is our objective.

2.3 Dynamic Resource Allocation

Thus, the company's decision making in both mobile and non-mobile settings can be treated as resource allocation in a dynamic environment, where "dynamic" means that the entire allocation plan of the company's stock develops concurrently with arrival of each new demand realization, revealed by a particular representative (see also [37, p. 151] terming dynamic environment in the vehicle routing context).

Dynamic resource allocation has been extensively studied by the CASTLE Laboratory at the Department of Operations Research and Financial Engineering of Princeton University, USA. The results have been reported among others in [32, 18, 13, 12, 10, 20]. The authors develop a framework for dealing with dynamic resource management problems, based on stochastic programming models and techniques. Their research was motivated initially by applications in freight transportation, such as truckload trucking and railroad car management, that can be described in short as managing reusable resources like containers, locomotives, drivers or crews, in a logistics network with customer orders arriving stochastically over time and requesting resources at certain locations to be occupied for certain activities. The problem of managing these resources efficiently is being complicated greatly by a dynamic nature of information input (e.g., customer orders) and a multitude of stochastic parameters including not only external ones like order quantities and locations, but also internal, like travel times, fuel consumption, costs, and others. The fact that the travel times and associated costs cannot be known with certainty for a movement decision to be made, and even that they may not become known until the movement is completed, makes our knowledge on availability of resources at any future time and at any location uncertain too. Since particular resources (like locomotives and crews) have to be coupled together to execute a decision, one have to be able to handle such resource "bundling" and "layering" properly. The framework is presented in the context of a railroad car distribution problem in the book chapter by Powell and Topaloglu [35].

The application of the above models and techniques is not restricted only to reusable resources in freight transportation, but can also be applied to distribution problems where a product is being distributed from plants to warehouses in anticipation of uncertain customer demand. Such application is demonstrated, for example, by Cheung and Powell in [12] and Godfrey and Powell in [20].

The research carried by the CASTLE Laboratory is presented well on its online resource [33].

The framework developed and successfully applied by the above authors to complex practical problems has rich modeling and solution capabilities. The problem we are studying here has much lower dimensionality and complexity, and can be fitted into the discussed framework. Though, we attempt to approach our problem "from scratch" because this should let us take a first closer look on it. Another argument for doing that are stochastic coefficients in our objective function, what was typically not assumed in the modeling of railroad car management as well as product distribution presented in [13, 12, 20]. Still, we believe that the methods developed in the context of the above resource management problems can be used to extend and investigate our setting, what we are going to approach in our further research.

2.4 Available to promise

The model that we are studying might be also coupled with the concept of *available-to-promise* (ATP), which can be defined as a functionality of responding to customer order requests by determination of supply/delivery quantities and their due dates. Since the availability of the product depends on many factors like production resources and operations scheduling, the ATP function links the customer orders with the enterprise resources, with the objective of improving responsiveness in order promising as well as reliability in order fulfillment [11, 5]. "The advanced ATP function is a decision-making mechanism that handles uncertainty and changes from external suppliers and customers, as well as from internal production processes. Its purpose is to improve profitability and customer service as well as to mitigate the discrepancy between forecast-driven push activities and order-driven pull activities across a supply chain system" [11]. The latter means that the ATP function resides typically on that point in the supply chain up to which the customer orders penetrate. Up from this point all downstream activities are order-driven, as opposed to the forecast-driven activities up to this point. So "the ATP system executes at the interface of the push-driven flow and pull-based flow systems" [5].

Keeping this in mind, let us take a look at our decision making from the following point of view. Let us still assume that the company distributes a final product to its customers. Let a new volume s_1 of the product be released and become available at the beginning of the coming week, according to the production schedule. Then the mission of the sales representatives would be to respond to the customer demand so that the available product (a limited resource) becomes allocated to the customers most efficiently, so that the overall customer satisfaction becomes as high as possible. If we assume that in the next week there will be the total product volume s_2 available, then we are interested again in a most efficient distribution of this amount by the representatives between the customers.

In other words, all together the representatives implement an ATP mechanism: they allocate, i.e. promise, some product quantities to the customers upon learning their individual demands, which become known over time. May the mobile technologies help the representatives to improve

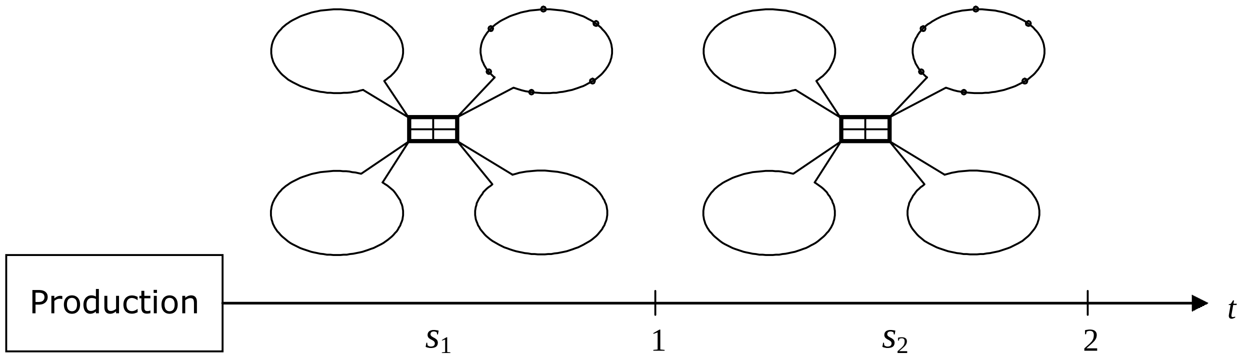


Figure 2: Distribution of the product in consecutive time periods

their performance, i.e. to improve the overall customer satisfaction, or, in other words, to improve the ATP mechanism? We are going to prove that yes, the representatives perform in this sense better when they utilize mobile technologies determining product quantities (note that we made initially in the section 1.1 the assumption that the company has weekly the same stock s just for evaluating its average performance given a stock level s). Also note that a representative may not meet an extraordinary large individual demand fully, even though the currently available stock would cover this demand: an ATP policy may prescribe to meet this demand only partially, because a particular stock must be kept for servicing the other customers. So each time an available-to-promise quantity is being determined according to some policy and allocated to the customer. Demands that cannot be covered fully in this week, will be scheduled to be met in the next week(s).

Now let us generalize the described operating mode in the following way. Let the customer orders penetrate the supply chain up to some particular point upstream. In other words, we consider now an assemble-to-order enterprise. There is no any finished product made to stock. Can we still utilize our model somehow? Yes, the model still may have sense: for example, in the following setting. Assume that the production of the enterprise is being assembled and sold according to contracts signed with the customers by the company's managers. These managers have to travel to the important customers and negotiate the contracts. For the production volumes secured by the contracts there will be needed particular enterprise resources, which are limited. Hence, these resources have to be efficiently allocated. Now not the sales representatives, but the managers are allocating limited resources to the customer orders. And these allocations seem to be more sophisticated to manage, since they involve planning of how the resources will be assigned for the assembling activities. Since all managers have to be aware of the available resources, and each has to be able to make a decision that would contribute to the company's overall performance, the mobile technologies may play here a more important role, either, than in the case of sales representatives, since any voice communication over phone makes less sense.

Note that everything what was said above for the sales representatives, applies to the managers, too, but at a higher degree of complexity. So our problem setting with sales representatives may find in fact a broader application. This decision making might be integrated also into a more complex environment like an ATP system managing a multi-commodity resource pool.

The rest of the paper is organised as follows. We proceed to the mathematical modeling of our approach in the next section. Section 4 demonstrates how does the model perform under a particularly simple demand distribution, and demonstrates some decision and evaluation techniques.

Section 5 generalizes them for any arbitrary dimensionality of a distribution system and any finite discrete distribution of customer demands, as well as presents some computational results.

3 Modeling

3.1 Servicing a single customer

Let us consider at first a system where a single representative provides service to a single customer. Let such configuration of a distribution system be called C|1|. We sketched it in Figure 3.

In C|1|, the company reserves at the beginning of the week s units of the good for meeting uncertain customer demand. The only representative, having received the whole stock s at his disposal, makes his way to the customer, learns his demand ξ_1, makes an allocation x_1 with the objective of maximizing this weeks's overall performance, and then makes his way back home. The single representative doesn't utilize any mobile facilities in his decision making (indeed, it wouldn't make much sense since he is the only one who makes allocations).

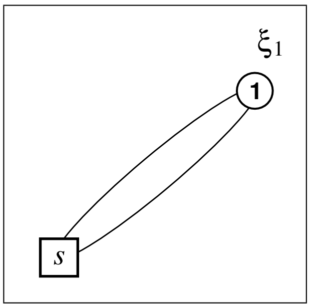


Figure 3: Configuration C|1|

Hence, the representative faces the following decision problem while making an allocation x_1:

$$\begin{aligned} \frac{x_1}{\xi_1} &\longrightarrow \max \\ x_1 &\leq s \\ x_1 &\leq \xi_1 \\ x_1 &\geq 0 \end{aligned} \tag{2}$$

where $\xi_1 = 0$ implies $x_1 = 0$ and the fraction in the objective equals 1 by definition. Note that (2) is a deterministic linear program since the decision is made upon learning the demand realization ξ_1.

Obviously, decision making in C|1| is particularly simple: the representative should allocate to the customer as much of the good as possible, maximizing by that the customer satisfaction. This resolves question Q1 in the given configuration.

Let us now discuss this distribution system in the context of the question Q3: what overall performance should we *expect* to attain in C|1| in any given week? Indeed, in presence of uncertain customer demand, we are interested in *expected performance*. The latter can also be understood as the average system's performance on the long run, given a weekly initial stock s. How can it be expressed?

Obviously, one should consider the expectation of the optimal objective value in (2). As one may see, the above problem incorporates random parameter ξ_1 varying from week to week. Hence, the optimal objective value in (2), being dependent on this random parameter, becomes a random variable, too. Let us rewrite (2) in the following form:

$$Z(\xi_1, s) \;=\; \max_{x_1} \left\{ \frac{x_1}{\xi_1} \middle| \, x_1 \le s,\; x_1 \le \xi_1,\; x_1 \ge 0 \right\} \tag{3}$$

where $Z(\xi_1, s)$ denotes the optimal objective value expressing the customer satisfaction in case of initial stock s and demand realization ξ_1. Hence, the expectation of the random variable $Z(\xi_1, s)$ with respect to the probability distribution of ξ_1 provides us with the expected performance in configuration C|1| given the initial stock s, which we define respectively as

$$E|1|(s) \;=\; E_{\xi_1}[Z(\xi_1, s)]. \tag{4}$$

Obviously, we cannot speak about applying mobile facilities for joint resource sharing in our distribution system with the only representative. The next section introduces a system of a bit higher complexity and illustrates our approach in modeling the effects of mobile technologies.

3.2 Servicing 2 customers

3.2.1 Non-mobile decision making

Let us consider now servicing of two customers in the following two different configurations:

C|1|1| : two representatives, each servicing just a single customer, with no mobile facilities, and

C|2| : a single representative, servicing the two customers in the sequence.

These two configurations are depicted in Figure 4. Note that in both cases there are no any mobile facilities in use.

Figure 4: Configurations C|1|1| and C|2|

In both cases the company reserves the same initial stock s at the beginning of the week. The random customer demands ξ_1 and ξ_2 are assumed to be independent and identically distributed.

Then, in C|1|1| the company allocates the same quota $a = s/2$ for meeting each customer's demand. The representatives will have to make such decisions x_1 and x_2 which maximize the overall performance:

$$
\begin{array}{llll}
\dfrac{x_1}{\xi_1} & + & \dfrac{x_2}{\xi_2} & \longrightarrow \quad \max \\
x_1 \leq a & & & \\
x_1 \leq \xi_1 & & & \\
x_1 \geq 0 & & & \\
& & x_2 \leq a & \\
& & x_2 \leq \xi_2 & \\
& & x_2 \geq 0 &
\end{array}
\tag{5}
$$

where $\xi_i = 0$ implies $x_i = 0$ and the corresponding fraction in the objective equals 1 by definition. Note that (5) is a deterministic linear program since both decisions are made upon learning the demand realizations ξ_1 and ξ_2.

Obviously, we can split (5) into two independent linear programs:

$$
\begin{array}{lll}
\dfrac{x_1}{\xi_1} \longrightarrow \max & \qquad\qquad & \dfrac{x_2}{\xi_2} \longrightarrow \max \\
x_1 \leq a & & x_2 \leq a \\
x_1 \leq \xi_1 & & x_2 \leq \xi_2 \\
x_1 \geq 0 & & x_2 \geq 0
\end{array}
\tag{6}
$$

which the representatives solve independently of each other upon arriving to their customers and learning their demands. Thus, the overall performance in C|1|1| sums up from the two individual performances and becomes maximized if (and only if) each representative maximizes his individual performance. We will refer further to the above two problems as to (6.1) and (6.2), respectively.

How should the allocation decisions be made in the configuration C|2|? In this case the same initial stock s has to be distributed by a single representative between the two customers sequentially and, again, in the most efficient way. "Sequentially" means here that the allocations of the good to the customers are being made one-by-one, so that the allocation to the 1st customer has to be made before the 2nd customer's demand becomes known. Then, arriving to the 1st customer, the representative faces the following optimization problem:

$$
\begin{array}{llll}
\dfrac{x_1}{\xi_1} & + & E_{\xi_2}\left[\max\limits_{x_2} \dfrac{x_2}{\xi_2}\right] & \longrightarrow \quad \max \\
x_1 \leq s & & & \\
x_1 \leq \xi_1 & & & \\
x_1 \geq 0 & & & \\
& & x_2 \leq s - x_1 & \\
& & x_2 \leq \xi_2 & \\
& & x_2 \geq 0 &
\end{array}
\tag{7}
$$

Since only the demand of the 1st customer is known to the representative at the time, he has to work on the overall performance in the sense of its expectation, which has to be maximized. Note that his 1st stage allocation x_1 leaves the space of as much as $s - x_1$ units for the allocation to be made on the 2nd stage. So the more we allocate to the 1st customer, the less we leave for the 2nd one and so the lower the expectation of the 2nd customer's satisfaction becomes, and vice versa. Of course, the 2nd-stage decision is supposed to utilize the available stock $s - x_1$ in a most efficient way, that's why the objective function in (7) incorporates the expected value of the highest possible 2nd-stage performance.

Problem (7) occurs to be a so called *two-stage stochastic linear program.* We can rewrite it also in the following form:

$$
\begin{aligned}
\frac{x_1}{\xi_1} \; + \; \mathcal{Q}(x_1) \quad &\longrightarrow \quad \max \\
x_1 \; &\leq \; s \\
x_1 \; &\leq \; \xi_1 \\
x_1 \; &\geq \; 0 \\
\mathcal{Q}(x_1) \; &= \; E\big[Q(x_1, \xi_2)\big] \\
Q(x_1, \xi_2) \; &= \; \max\left\{ \frac{x_2}{\xi_2} \;\middle|\; x_2 \leq s - x_1, \; x_2 \leq \xi_2, \; x_2 \geq 0 \right\}
\end{aligned}
\tag{8}
$$

As one may see, x_1 is the only decision variable in this problem. Function $\mathcal{Q}(x_1)$, called the *expected recourse function*, expresses the expectation of the *recourse function* $Q(x_1, \xi_2)$, which is itself the optimal objective value of the following embedded *recourse problem*:

$$
\begin{aligned}
\frac{x_2}{\xi_2} \; &\longrightarrow \; \max \\
x_2 \; &\leq \; s - x_1 \\
x_2 \; &\leq \; \xi_2 \\
x_2 \; &\geq \; 0
\end{aligned}
\tag{9}
$$

Hence, a recourse program is the one "in which some decisions or recourse actions can be taken after uncertainty is disclosed" [8, p. 52]. Note that this is just a deterministic linear program and is exactly of the same type as problem (2) faced by the single representative in configuration C|1|. Its optimal objective value $Q(x_1, \xi_2)$ represents in our case the 2nd customer's satisfaction which the representative would deliver facing the demand realization ξ_2 and having $s - x_1$ units of the good at his disposal. Hence, $\mathcal{Q}(x_1)$ is the expected satisfaction of the 2nd customer given that we allocate x_1 units to the 1st one, i.e. that we leave $s - x_1$ units for meeting the 2nd customer's demand.

Note that in order to evaluate $Q(x_1, \xi_2)$ for any given x_1 we have to solve each time the embedded linear program. In order to evaluate $\mathcal{Q}(x_1)$ for any given x_1 we have to compute the expectation of the optimal objective value of this embedded linear program, with respect to the probability distribution of ξ_2. See the textbooks [23, p. 16], [24, pp. 9–11, 25–26], [8, p. 54], [43, pp. 11–12] for more details on two-stage stochastic programs.

After a decision x_1 is fixed, the representative makes his way to the 2nd customer, learns his demand ξ_2 and makes an allocation x_2 by solving the second stage problem (9) where x_2 is the only variable and all parameters are deterministically known. One may also see that this is the same optimization problem as (6.2) where $s - x_1$ is substituted for a.

Note that by solving the optimization problems (6) and (8),(9) in configurations C|1|1| and C|2|, respectively, we resolve the question Q1 for them both!

Let us examine now these two configurations in the context of the question Q3: which of them should we *expect* to deliver a better overall performance? Indeed, in presence of uncertain customer demands, we are interested in *expected overall performance* attained in these two configurations.

To this end, let us refer to the problems (5) and (7). They both incorporate parameter(s) ξ_i varying randomly from week to week and, hence, their optimal objective values are dependent on the random parameter(s) and become by itself random variables, too, which we can denote by $Z|1|1|$ and $Z|2|$, respectively. Further, let us denote their expectations by $E|1|1|$ and $E|2|$, respectively. As one can see, the latter two represent the expected overall customer satisfaction in the configurations C|1|1| and C|2|.[a] Then, important for us is to notice that the constraints $x_1, x_2 \leq s/2$ in (5) are relaxed to $x_1 + x_2 \leq s$ in (7). Hence, $E|2|$ is computed over a broader set of implementable decisions, and we can conclude:

$$E|1|1| \;\leq\; E|2|. \tag{10}$$

That is, we should expect an overall performance in C|2| to be at least not worse than that in C|1|1|.[b] Our objective remains then to quantify this advantage.

How is this all related with mobile technologies? Or, in other words, how can the above comparison of the scenarios C|1|1| and C|2| help us to investigate the effects of mobile technologies? We proceed now to the basic idea behind our approach.

3.2.2 Mobile decision making

Let us introduce the following configuration C$\mathcal{M}$|1|1| for our distribution system: the same two representatives as in the case C|1|1| now utilize mobile technologies in their decision making (as described in section 1.2). The transition from C|1|1| to the new *mobile* configuration C$\mathcal{M}$|1|1| is depicted in Figure 5.

We interpret the diagram to the right as follows: the two representatives are visiting their customers as they did, but now equipped with mobile devices. The same initial stock s is not divided between the representatives anymore. Instead, they make allocations from the common stock according to the mobile policy formulated in section 1.2.

Let us assume without loss of generality that the 1st representative arrives to his customer before the 2nd does. He makes then the following steps in his decision making:

1. learning the customer's demand;

[a] For rigorous proof see Proposition 1 in section 3.4.
[b] For rigorous proof see Proposition 2 in section 3.4.

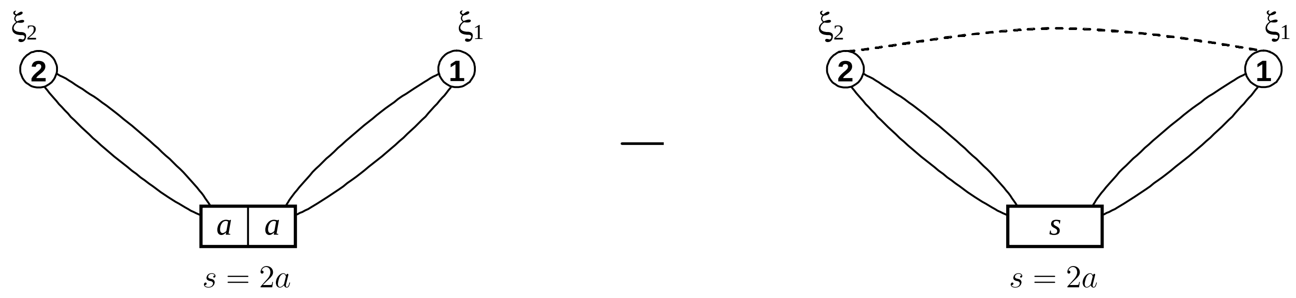


Figure 5: Configurations C|1|1| and C$\mathcal{M}$|1|1|

2. retrieving the information from the server on the currently available stock (=s) and the number of customers not yet served (=2);
3. making an allocation that would maximize the overall system's performance;
   note that without knowing the demand realizations of the pending customers, he is going to maximize the *expected* overall performance;
4. uploading the decision to the central computer.

We can conclude thereby that deciding on the allocation x_1, he faces exactly the same decision problem as the single representative in C|2| does as he arrives to his 1st customer. So they both solve the same optimization problem (8).

Let us keep tracing the case C$\mathcal{M}$|1|1|: the 1st representative, having made an allocation x_1, uploads this information from his mobile device to the central computer. The 2nd representative will have then as much room as $s - x_1$ units of the good for making an allocation x_2, what he does by solving a decision model that occurs to be the problem (9) which is also being solved by the single representative in the scenario C|2| upon arriving to the 2nd customer.

Thus, decision making in C$\mathcal{M}$|1|1| is identical with that in C|2|. With that, we resolve question Q2 for the case C$\mathcal{M}$|1|1|! As we see, the two representatives equipped with mobile facilities imitate fully the work of a single representative servicing the two customers sequentially. We have depicted this fact in the Figure 5 by the dashed line connecting the customers: it traces the route of such imaginary ("virtual") single representative.

Another conclusion that we can make here is the following one: by solving the same optimization problems as the single representative solves while making allocations to the same customers, the two mobile representatives obviously deliver on the long run the same average overall performance $E\mathcal{M}|1|1|$ as the single representative does, and we have the following identity:

$$E\mathcal{M}|1|1| \;=\; E|2|. \tag{11}$$

At the same time, we have shown above with the formula (10) that the single representative in configuration C|2| outperforms the two separate (non-mobile) representatives in C|1|1|. Hence, the above identity proves also the advantage of the mobile solution against the non-mobile one, since it empowers the two representatives to achieve the performance of the single "virtual" one, and we come to the following relationship between the expected overall performances:

$$E|1|1| \;\leq\; E\mathcal{M}|1|1| \;=\; E|2|.$$

Let us make now the following remark: the optimal objective values in all above decision models and their expected values are dependent on the initial weekly stock s, since s is incorporated as a parameter in all these models, and we would like to rewrite the above relationship in the following form:

$$E|1|1|(s) \;\leq\; E\mathcal{M}|1|1|(s) \;=\; E|2|(s). \tag{12}$$

Our main objective remains then the quantification of such advantage delivered by mobile decision making, for any initial stock s.

Note that the expected overall performance in the non-mobile case C$|1|1|$ expresses as the sum of the expected individual performances of the two representatives, what provides us with:

$$E|1|1|(s) \;=\; E|1|(a) \;+\; E|1|(a) \;=\; 2 \cdot E|1|(s/2). \tag{13}$$

Expressing $E|2|$ requires more computational efforts, though — as we will see later.

The above illustrates our approach in studying the economic effects of mobile technologies. Let us now briefly extend this modeling for the case of servicing 3 customers, after which we generalize our approach for any configuration of a distribution system.

3.3 Servicing 3 customers

3.3.1 Servicing 3 customers by 3 representatives

Let us consider here the following two configurations:

C$|1|1|1|$ $\equiv$ C$|1|^3$: *non-mobile case* : 3 representatives, each assigned a single customer;

C$\mathcal{M}|1|1|1|$ $\equiv$ C$\mathcal{M}|1|^3$: *mobile case* : the same 3 representatives, utilizing mobile facilities.

Figure 6 sketches these two configurations.

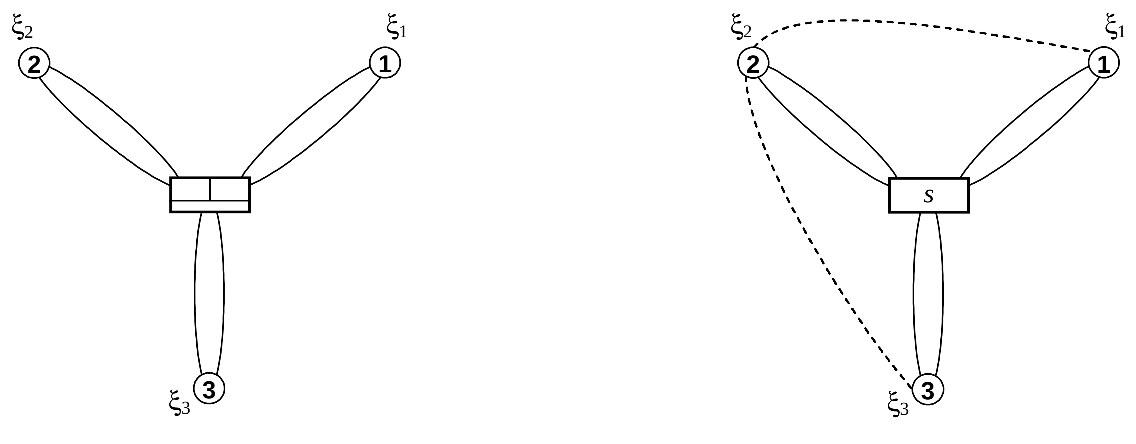


Figure 6: Configurations C$|1|^3$ and C$\mathcal{M}|1|^3$

First of all, let us answer the questions Q1 and Q2 in these two configurations, respectively: how should decisions in them both be made given the weekly initial stock s? We keep assuming the customer demands to be independent and identically distributed.

Our approach here will be the same as in the previous section. Let us refer to the case $\mathrm{C}|\mathsf{1}|^3$ first. Obviously, the company divides the initial stock between the representatives equally so that each of them receives the quota $a = s/3$ for meeting his customer's demand. After that they make their decisions independently of each other upon arriving to the assigned customers and learning their demands. That is, similar to (6), the representatives face the following decision problems:

$$
\begin{array}{lll}
\dfrac{x_1}{\xi_1} \longrightarrow \max & \dfrac{x_2}{\xi_2} \longrightarrow \max & \dfrac{x_3}{\xi_3} \longrightarrow \max \\
\quad x_1 \leq a & \quad x_2 \leq a & \quad x_3 \leq a \\
\quad x_1 \leq \xi_1 & \quad x_2 \leq \xi_2 & \quad x_3 \leq \xi_3 \\
\quad x_1 \geq 0 & \quad x_2 \geq 0 & \quad x_3 \geq 0
\end{array}
\tag{14}
$$

The overall performance in any given week sums up from individual performances of the representatives which turn out to be the optimal objective values of the above three problems.

Let us now refer to the case $\mathrm{C}\mathcal{M}|\mathsf{1}|^3$. The initial stock s is not divided between the representatives anymore. We assume here without loss of generality that the representatives visit their customers and make their allocations from the common stock s in the timely sequence $1 \rightarrow 2 \rightarrow 3$, as the dashed line in Figure 6 shows. How does the 1st representative decide on the allocation for the 1st customer upon learning his demand?

Remember the reasoning which gave us the answer to this question for the mobile scenario $\mathrm{C}\mathcal{M}|\mathsf{1}|\mathsf{1}|$ in section 3.2.2: the representatives, utilizing mobile facilities, together imitate the work of a single "virtual" representative who visits the customers in the same sequence as they are being visited by the "real" representatives. Thus the 1st "real" representative faces upon arriving to the 1st customer the same problem as the single "virtual" representative would do. This was found in $\mathrm{C}\mathcal{M}|\mathsf{1}|\mathsf{1}|$ to be the problem (8).

Using the same logic now for the case of having 3 customers, we make the 1st representative solve the following optimization problem in scenario $\mathrm{C}\mathcal{M}|\mathsf{1}|^3$:

$$
\begin{array}{l}
\dfrac{x_1}{\xi_1} + E\left[\max\limits_{x_2}\left\{\dfrac{x_2}{\xi_2} + E\left[\max\limits_{x_3} \dfrac{x_3}{\xi_3}\right]\right\}\right] \quad \longrightarrow \quad \max \\
x_1 \leq s \\
x_1 \leq \xi_1 \\
x_1 \geq 0 \\
\qquad\qquad x_2 \leq s - x_1 \\
\qquad\qquad x_2 \leq \xi_2 \\
\qquad\qquad x_2 \geq 0 \\
\qquad\qquad\qquad\qquad x_3 \leq s - x_1 - x_2 \\
\qquad\qquad\qquad\qquad x_3 \leq \xi_3 \\
\qquad\qquad\qquad\qquad x_3 \geq 0
\end{array}
\tag{15}
$$

Here the 1st representative arranges his decision x_1 to be such that his individual performance plus the expected overall performance of his colleagues gets maximal. Since the information on the 2nd and 3rd customers' demand is not available yet, this is the only way for him to contribute to the maximization of the overall performance. On the 2nd stage, the decision x_2 to be made by the 2nd representative will have to allocate so much from the available stock $s - x_1$ that his individual performance plus the expected performance of the 3rd colleague gets maximal. The 3rd representative will maximize just his individual performance since his allocation is the last to be, and he has as much room for it as $s - x_1 - x_2$ units.

As one may see, the above problem is a *3-stage stochastic linear program.* The only decision variable here is x_1, since we cannot decide on the other allocations x_2 and x_3 before the corresponding demand realizations become known. The problem can also be rewritten in the following form:

$$
\begin{aligned}
&\frac{x_1}{\xi_1} + Q(x_1) \quad \longrightarrow \quad \max \\
&\quad x_1 \le s \\
&\quad x_1 \le \xi_1 \\
&\quad x_1 \ge 0 \\
&Q(x_1) = E\big[Q(x_1, \xi_2)\big] \\
&Q(x_1, \xi_2) = \max_{x_2} \left\{ \frac{x_2}{\xi_2} + E\left[\max \frac{x_3}{\xi_3}\right] \middle| \begin{array}{lll} x_2 \le s - x_1, & x_2 \le \xi_2, & x_2 \ge 0 \\ x_3 \le s - x_1 - x_2, & x_3 \le \xi_3, & x_3 \ge 0 \end{array} \right\}
\end{aligned}
\tag{16}
$$

Again, $Q(x_1)$ is the *expected recourse function* expressing the expected value of the *recourse function* $Q(x_1, \xi_2)$. The latter is itself the optimal objective value of the embedded 2-stage stochastic program that delivers an optimal decision of the 2nd representative under demand realization ξ_2. This 2-stage program, in turn, embeds another linear program, which is single-stage and corresponds to the final decision to be made by the 3rd representative.

So, we may say that the 1st representative, in presence of uncertainty concerning the 2nd and 3rd customers' demands, fixes such an allocation x_1 that would maximize this week's overall performance, by solving problem (16).

After that, the 2nd representative will make a decision x_2 upon learning the demand ξ_2. Note that there are $s - x_1$ units of the good left for 2 sequential allocations — to the customers 2 and 3. Then, how should the decision x_2 be made? In fact, we have discussed this already in the previous section: problem (8) gives us an optimal decision for the 1st of two mobile representatives provided with an initial stock s. Hence, it fits our case, too: the 2nd representative should be treated as the first one, the 3rd — as the second one, and their initial stock occurs to be $s - x_1$.

So, we may say that our 2nd representative, in presence of uncertainty concerning the 3rd customer's demand, fixes such an allocation x_2 that would maximize his and his 3rd colleague's overall performance, by adjusting properly problem (8) and solving it.

Finally, the 3rd representative will have $s - x_1 - x_2$ units left after the decisions x_1 and x_2 have been made. He is the only one to make use of this remaining stock. Then, how should the decision x_3 be made? In fact, we have discussed this question already in configuration C|1|: problem (2) gives us an optimal decision for the single representative provided with an initial stock s. Hence, it fits our case, too: the 3nd representative should be treated as this representative, and his initial stock occurs to be $s - x_1 - x_2$.

So, we may say that the 3rd representative fixes such an allocation x_3 that would maximize his individual performance, by adjusting properly problem (2) and solving it.

With that we resolve question Q2 in configuration $\mathrm{C}\mathcal{M}|1|^3$. We would like to summarize the above as follows: three representatives utilizing mobile facilities, jointly execute sequential resource allocations to $N = 3$ customers.

We can now turn to the question Q3 and compare the expected overall performance of the system in the configurations $\mathrm{C}|1|^3$ and $\mathrm{C}\mathcal{M}|1|^3$. We can use the same reasoning here as we used for configurations C|1|1| and C|2|in section 3.2.1. Namely, let us refer to the decision problems (14) and (15). Let $E|1|^3$ and $E\mathcal{M}|1|^3$, respectively, denote the expectations of their optimal objective values (which are by itself random variables, being dependent on random parameter(s)). Further, $E|1|^3$ and $E\mathcal{M}|1|^3$ represent the expected overall customer satisfaction[c] in the configurations $\mathrm{C}|1|^3$ and $\mathrm{C}\mathcal{M}|1|^3$, and important for us becomes to notice that the constraints $x_1, x_2, x_3 \leq s/3$ in (14) are relaxed to $x_1 + x_2 + x_3 \leq s$ in (15). Hence, $E\mathcal{M}|1|^3$ is computed over a broader set of implementable decisions, and we can conclude:

$$E|1|^3(s) \;\leq\; E\mathcal{M}|1|^3(s). \tag{17}$$

That is, we should expect an overall performance in $\mathrm{C}\mathcal{M}|1|^3$ to be at least not worse than that in $\mathrm{C}|1|^3$, for any given initial stock s.[d] Our objective remains then to quantify this advantage.

3.3.2 Servicing 3 customers by 2 representatives

Let us consider now the following two configurations:

C|2|1| : *non-mobile case* : the 1st representative services the customers 1 and 3 sequentially, while the 2nd representative is assigned the only customer no. 2;

$\mathrm{C}\mathcal{M}|2|1|$: *mobile case* : the same 2 representatives, utilizing mobile facilities.

Figure 7 depicts these two configurations. As always, we assume the customer demands to be independent and identically distributed.

In C|2|1|, the company divides the initial stock s between two representatives. Consequently, the 1st representative receives twice as more as his colleague does: we can represent their quotas as $2a$ and a, respectively, with $a = s/3$. Then, keeping in mind that the representatives work separately of each other, we already know at this point how are the allocation decisions in this configuration be made: the decision making of a single representative servicing 2 customers in a sequence was discussed by us in configuration C|2|. Respectively, the decision making of a single representative servicing a single customer was discussed in configuration C|1|.

[c]For rigorous proof see Proposition 1 in section 3.4.
[d]For rigorous proof see Proposition 2 in section 3.4.

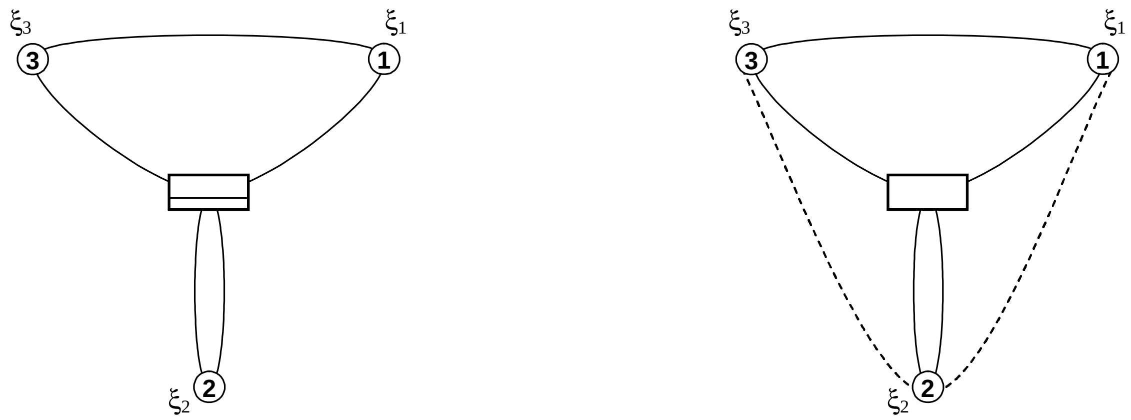


Figure 7: Configurations C|2|1| and C$\mathcal{M}$|2|1|

In C$\mathcal{M}$|2|1|, the initial stock is not divided between the representatives, so that they make their allocations from the common resource pool. We assume again without loss of generality that the representatives visit their customers and make their allocations in the timely sequence $1 \to 2 \to 3$, as the dashed line in Figure 7 shows. We can easily see that sequence of events in this mobile environment copies that of the system C$\mathcal{M}$|1|^{3}discussed in the previous section. In other words, here two representatives utilizing mobile facilities, jointly execute sequential resource allocations to $N = 3$ customers. Hence, modeling carried out for C$\mathcal{M}$|1|^3 applies fully to the current case C$\mathcal{M}$|2|1|.

Thus we answered with the above questions Q1 and Q2. Let us turn now to the question Q3. As one may guess, we should expect the mobile solution C$\mathcal{M}$|2|1| to deliver a better (at least a not worse) performance than its non-mobile counterpart C|2|1|. Indeed, the expected overall performance in the non-mobile case is the sum of the expected individual performances of the two representatives, so that

$$E|2|1|(s) \;=\; E|2|(2a) \;+\; E|1|(a) \tag{18}$$

where $E|2|$ is being computed as the expectation of the optimal objective value in (7), and $E|1|$ — as the expectation of the optimal objective value in (2), which enforce in our case the constraints $x_1 + x_3 \leq 2a$ and $x_2 \leq a$, respectively.

The expected overall performance in C$\mathcal{M}$|2|1| coincides with that in C$\mathcal{M}$|1|^3:

$$E\mathcal{M}|2|1|(s) \;=\; E\mathcal{M}|1|^3(s), \tag{19}$$

which is being computed as the expectation of the optimal objective value in (15), where any sequence of decisions must satisfy $x_1 + x_2 + x_3 \leq s$, what provides us with a broader set of implementable decisions, and hence, a better (a not worse) expected overall performance in the mobile case C$\mathcal{M}$|2|1| comparing to C|2|1|, again.

The above illustrates our approach in modeling the economic effects of mobile technologies. The next section generalizes our modeling for the case of any configuration of a distribution system.

3.4 General case: any configuration

The ideas presented in the preceding sections stay valid for any other topology of a distribution system, as Figure 8 demonstrates.

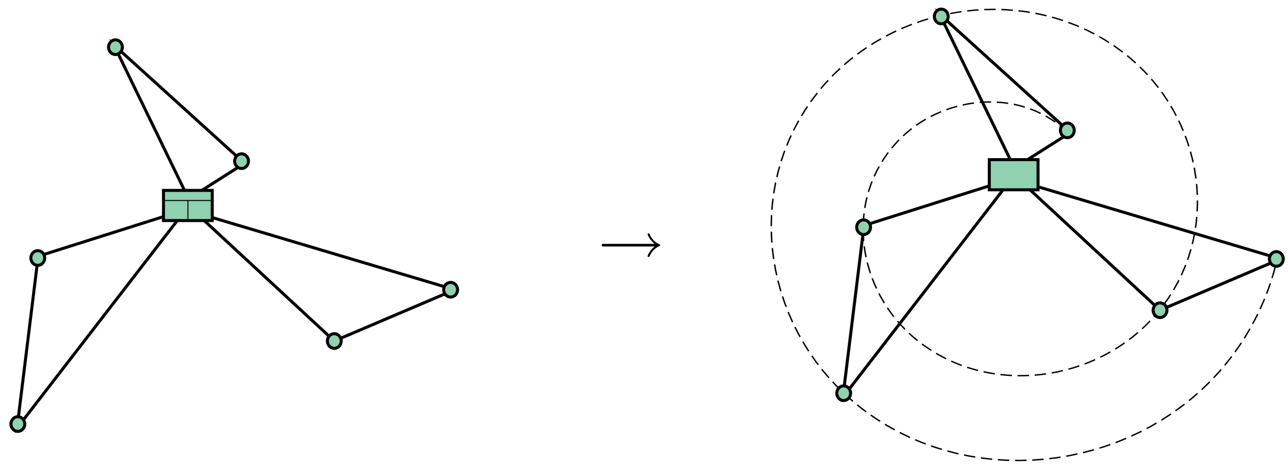


Figure 8: Transition to the mobile decision making

The picture to the left presents there a distribution system where 3 representatives provide service to 2 customers each, while the mobile solution is not implemented. The total stock is being equally divided between the representatives (assuming the customer demands to follow the same probability distribution).

The picture to the right illustrates the transition to the mobile solution in this system. The same representatives provide their service to the same customers still, so that the assignment of the customers to the agents is kept. At the same time, the total stock is not being divided anymore. The allocations are made from the common resource pool using centralized data on the currently available stock and the number of pending customers, with the objective of maximizing the overall performance. If we assume that the customers are being visited by the representatives in the timely sequence (shown in the figure by the dashed line), we conclude that the overall system's performance equals the performance of an imaginary single representative visiting the customers in the same sequence. He, in turn, is assumed to outperform several independent representatives in a non-mobile scenario.

Let us now formalize our approach. We consider a system with N customers which are partitioned into K groups of $n_1, \ldots, n_K$ members, with each group assigned to a sales representative. The customers have individual weekly demands $\xi_1, \ldots, \xi_N$ which are not known in advance. At the beginning of the week the company divides its initial stock s between the representatives who recieve their quotas $s_1, \ldots, s_K$ according to the expected customer demands in their groups. Then, each representative executes sequential product allocations to the assigned customers, upon learning each individual demand. Each representative makes his allocation decisions with the objective of maximizing his individual performance while not exceeding his individual stock. We assume customer demands to be independent and identically distributed.

We denote this (non-mobile) configuration of the distribution system by $\mathrm{C}|n_1|\dots|n_K|$.

Now let us introduce mobile facilities into the company's decision making and denote this new (mobile) configuration by $\mathrm{C}\mathcal{M}|n_1|\dots|n_K|$. The company doesn't divide the stock s between the representatives anymore but rather uses it as a joint resource pool for centralized product allocations, by utilizing mobile facilities. In this configuration the representatives imitate together the work of a single imaginary ("virtual") representative who services all N customers in the same timely sequence as they are being serviced by the "real" representatives. Let us denote such equivalent "virtual" configuration by $\mathrm{C}|N|$. We will further assume without loss of generality the above servicing sequence to be $1,\dots,N$.

Hence, the decision making in both cases concerns efficient resource allocations in a sequence of n customers, where $n = n_1,\dots,n_K$ for each single representative in the 1st case, and $n = N$ in the 2nd case.

Let now s denote the initial stock (initial resource supply) in a sequence of n customers. Then the very first decision x_1 in this sequence, being determined upon learning the demand ξ_1, expresses as an optimal solution to the following problem:

$$
\begin{array}{llll}
\dfrac{x_1}{\xi_1} + E_{\xi_2}\left[\max\limits_{x_2}\left\{\dfrac{x_2}{\xi_2} + E_{\xi_3}\left[\ \dots + E_{\xi_n}\left[\max\limits_{x_n}\dfrac{x_n}{\xi_n}\right]\right]\right\}\right] & & & \longrightarrow \quad \max \\
x_1 \le s & & & \\
x_1 \le \xi_1 & & & \\
x_1 \ge 0 & & & \\
 & x_2 \le s - x_1 & & \\
 & x_2 \le \xi_2 & & \\
 & x_2 \ge 0 & & \\
 & & \ddots & \\
 & & & x_n \le s - \sum_{i=1}^{n-1} x_i \\
 & & & x_n \le \xi_n \\
 & & & x_n \ge 0
\end{array}
\tag{20}
$$

where, as always, $\xi_i = 0$ implies $x_i = 0$ and the corresponding fraction in the objective equals 1 per definition.

Problem (20) is a *n-stage stochastic linear program.* It embeds by itself another stochastic program corresponding to the allocation x_2 to be done in the 2nd stage upon learning the demand ξ_2. The latter, in turn, embeds the stochastic program corresponding to subsequent decision x_3, and so on, up until the final n-th stage represented by a deterministic program delivering an optimal decision x_n under the demand realization ξ_n. At each of the stages we take into consideration the expectation of the embedded program with respect to the distribution of demand to be disclosed in the next stage. In other words, problem (20) embeds a sequence of nested stochastic programs.

We can give an equivalent formulation of (20) in terms of a dynamic program [8, p. 128]. Let us consider decision making on the k-th stage ($k = 1,\dots,n$). Upon disclosing a customer demand ξ_k there should an optimal decision x_k be made which depends on ξ_k and also on the entire history of the preceding decisions $x_1,\dots,x_{k-1}$. Let $\tilde{s}$ denote the stock available at the beginning of the current stage. We obtain then x_k as an optimal solution to the following problem:

$$
\begin{aligned}
Z_{n-k+1}(\xi_k, \tilde{s}) \;&=\; \max \; \frac{x_k}{\xi_k} \;+\; E\left[Q_{n-k}(x_k, \xi_{k+1})\right] \\
x_k \;&\leq\; \tilde{s} \\
x_k \;&\leq\; \xi_k \\
x_k \;&\geq\; 0
\end{aligned}
\tag{21}
$$

where

$$
Q_{n-k}(x_k, \xi_{k+1}) \;=\; Z_{n-k}(\xi_{k+1}, \tilde{s} - x_k)
$$

with

$$
Q_0 \equiv 0
$$

The subscript $n-k+1$ of the optimal objective value $Z_{n-k+1}(\xi_k, \tilde{s})$ corresponds to the number of customers up from the current one to the end of the sequence. Function $Q_{n-k}(x_k, \xi_{k+1})$ is called the *recourse function*. It represents the optimal objective value of the embedded program if the demand realization in the next stage will occur to be ξ_{k+1} while the available stock will amount to $\tilde{s} - x_k$. The expected value of $Q_{n-k}(x_k, \xi_{k+1})$ defines the *expected recourse function*:

$$
\mathcal{Q}_{n-k}(x_k) \;=\; E_{\xi_{k+1}}\left[Q_{n-k}(x_k, \xi_{k+1})\right], \tag{22}
$$

which, hence, expresses the expected performance in all the future stages given a decision in the current stage.

Thus, the first-stage decision x_1 turns out to be a solution to the following problem:

$$
\begin{aligned}
Z_n(\xi_1, s) \;&=\; \max \; \frac{x_1}{\xi_1} \;+\; \mathcal{Q}_{n-1}(x_1) \\
x_1 \;&\leq\; s \\
x_1 \;&\leq\; \xi_1 \\
x_1 \;&\geq\; 0
\end{aligned}
\tag{23}
$$

All subsequent stages $k = 2, \ldots, n$ are absorbed in (23) into the function $\mathcal{Q}_{n-1}(x_1)$ through the corresponding expected values [43, p. 24].

Let us summarize: our ability to solve the sequence of problems (21) for any given s and any number of stages n would resolve the questions Q1 and Q2 in the non-mobile and mobile configurations, respectively.

Let us now turn to the question Q3: which configuration of them two should we expect to deliver a better overall performance, given the same initial stock s?

Let us at first clarify how do we define the *expected overall performance*. We introduced in section 1 the measure of the company's overall performance in any given week as

$$
G(\mathbf{x}|\xi) \;=\; \frac{x_1}{\xi_1} \;+\; \ldots \;+\; \frac{x_N}{\xi_N}\;, \tag{24}
$$

which has to be computed ex-post — i.e. after all allocations in response to the customer demands have been committed. Hence, we define the expectation of $G(\mathbf{x}|\xi)$ with respect to the distribution of ξ as the system's expected overall performance:

$$E_\xi \Big[G(\mathbf{x}(\xi)|\xi) \Big], \tag{25}$$

where the notation $\mathbf{x}(\xi)$ emphasizes the fact that the decision vector $\mathbf{x}$ is being computed as a reaction to the customer demand ξ, according to some decision policy.

Let us now consider a single representative in the non-mobile environment, who services n customers sequentially given an initial stock s. We define his expected individual performance exactly in the same way as in (24)–(25) — just by substituting n for N — and denote it by $E|n|$. Let us now show that if he makes his sequential allocation decisions by solving the sequence of problems (21) ($k = 1, \ldots, n$) then his expected individual performance equals the expectation of the optimal objective value of his first-stage problem (23), i.e. there holds the following

Proposition 1 *(Expected overall performance)*

$$E|n|(s) \;=\; E_{\xi_1} \big[Z_n(\xi_1, s) \big]$$

Proof: Obviously, the above equality holds for $n = 1$. We will show that it holds for $n = 2$; this proof can be then easily extended for any other positive n.

With $n = 2$ customers in the sequence, the 1st allocation is being determined as an optimal solution to the problem (23) upon the available information ξ_1 and hence, is computed as $x_1^*(\xi_1)$. The 2nd allocation is determined as an optimal solution to (21) upon disclosing ξ_2 and depends on both $x_1^*(\xi_1)$ and ξ_2 — hence, it is computed as $x_2^*(\xi_1, \xi_2)$. According to the definition, we have:

$$\begin{aligned}
E|n|(s) \;&= E_{\xi_1,\xi_2} \left[\frac{x_1^*(\xi_1)}{\xi_1} \;+\; \frac{x_2^*(\xi_1,\xi_2)}{\xi_2} \right] \;=\; E_{\xi_1,\xi_2} \left[\frac{x_1^*(\xi_1)}{\xi_1} \right] \;+\; E_{\xi_1,\xi_2} \left[\frac{x_2^*(\xi_1,\xi_2)}{\xi_2} \right] \;= \\
&= E_{\xi_1} \left[\frac{x_1^*(\xi_1)}{\xi_1} \right] \;+\; E_{\xi_1} \left[E_{\xi_2} \left[\frac{x_2^*(\xi_1,\xi_2)}{\xi_2} \right] \right] \;=\; E_{\xi_1} \left[\frac{x_1^*(\xi_1)}{\xi_1} \;+\; E_{\xi_2} \left[\frac{x_2^*(\xi_1,\xi_2)}{\xi_2} \right] \right] \;= \\
&= E_{\xi_1} \big[Z_1(\xi_1, s) \big].
\end{aligned}$$

□

Hence, in order to express the expected performance of any sequential allocation decision making we can just consider the expected value of the corresponding first-stage program.

Let us now turn to the non-mobile configuration $\mathrm{C}|n_1| \ldots |n_K|$ and denote its expected overall performance by $E|n_1| \ldots |n_K|$. Having K separate representatives who service $n_1, \ldots, n_K$ customers while provided with $s_1, \ldots, s_K$ units of the good, we conclude that

$$E|n_1| \ldots |n_K|(s) \;=\; E|n_1|(s_1) \;+\; \ldots \;+\; E|n_K|(s_K). \tag{26}$$

In the mobile configuration $\mathrm{C}\mathcal{M}|n_1| \ldots |n_K|$ the entire decision making corresponds to that of a single "virtual" representative who services N customers sequentially, and, hence, we have

$$E\mathcal{M}|n_1|\dots|n_K|(s) \;=\; E|N|(s). \tag{27}$$

Our objective for the rest of this section becomes to show the advantage of mobile decision making, i.e., that the inequality $E|n_1|\dots|n_K|(s) \le E\mathcal{M}|n_1|\dots|n_K|(s)$ holds. Before we proceed to this matter, let us consider our n-stage stochastic linear program once again and introduce one more formulation of it, additionally to the nested formulation (20) and dynamic programming formulation (21). First of all, let us refer to the case $n = 2$, which we discussed already in section 3.2 in the context of configurations C|2| and C$\mathcal{M}$|1|1|. Problem (8) (rewritten here below to the left) corresponds there to the 1st decision in the sequence. According to Ruszczyński and Shapiro [43, pp. 16–19], we can formulate it also in the form shown here to the right:

$$\begin{array}{ll}
\max\limits_{x_1}\left\{\dfrac{x_1}{\xi_1} + E_{\xi_2}\left[\max\limits_{x_2}\dfrac{x_2}{\xi_2}\right]\right\} & \qquad \max\limits_{\substack{x_1\in\mathbf{R}\\ x_2(\cdot)\in\mathcal{X}}} E_\omega\left[\dfrac{x_1}{\xi_1} + \dfrac{x_2(\omega)}{\xi_2(\omega)}\right] \\
\quad x_1 \le s & \qquad\quad x_1 \le s \\
\quad x_1 \le \xi_1 & \qquad\quad x_1 \le \xi_1 \\
\quad x_1 \ge 0 & \qquad\quad x_1 \ge 0 \\
\qquad\qquad x_2 \le s - x_1 & \qquad\qquad\quad x_2(\omega) \le s - x_1 \\
\qquad\qquad x_2 \le \xi_2 & \qquad\qquad\quad x_2(\omega) \le \xi_2(\omega) \\
\qquad\qquad x_2 \ge 0 & \qquad\qquad\quad x_2(\omega) \ge 0
\end{array}$$

In this equivalent formulation, x_2 is not a second-stage variable anymore but a random variable $x_2(\cdot)$ from the space $\mathcal{X}$ of measurable functions from Ω to $\mathbf{R}$. There is, however, an important difference between random variables $\xi_2(\omega)$ and $x_2(\omega)$: the former represents the random data of the problem with a given distribution, while the latter — the recourse action whose distribution is not given [43, p. 19]. Hence, in this equivalent formulation we perform maximization of the expected value over all possible values of the decision variable x_1 and measurable functions $x_2(\cdot)$ that satisfy jointly the constraints.

This formulation can be generalized for the case of a n-stage problem. Considering the nested formulation (20), we come to the following equivalent of it [43, p. 36]:

$$\begin{array}{l}
\max\limits_{\mathbf{x}(\cdot)} E_\omega\left[\dfrac{x_1(\omega)}{\xi_1(\omega)} + \dfrac{x_2(\omega)}{\xi_2(\omega)} + \dots + \dfrac{x_n(\omega)}{\xi_n(\omega)}\right] \\
\quad x_1(\omega) \le s \\
\quad x_1(\omega) \le \xi_1(\omega) \\
\quad x_1(\omega) \ge 0 \\
\qquad\quad x_2(\omega) \le s - x_1(\omega) \\
\qquad\quad x_2(\omega) \le \xi_2(\omega) \\
\qquad\quad x_2(\omega) \ge 0 \\
\qquad\qquad\qquad \ddots \\
\qquad\qquad\qquad\quad x_n(\omega) \le s - \sum_{i=1}^{n-1} x_i(\omega) \\
\qquad\qquad\qquad\quad x_n(\omega) \le \xi_n(\omega) \\
\qquad\qquad\qquad\quad x_n(\omega) \ge 0 \\
\quad x_i(\omega) = E_\omega\big[x_i(\omega) \mid \xi_{[1,i]}(\omega)\big], \quad i = 1, \dots, n
\end{array} \tag{28}$$

It is necessary to make the following notes here:

- In the above problem, both ξ_1 and x_1 are deterministic. But, for notational convenience, we will treat them as random variables $\xi_1(\omega)$ and $x_1(\omega)$ with some constant (deterministically known) values $\xi_1(\omega) \equiv \xi_1$ and $x_1(\omega) \equiv x_1$, respectively.
- $\xi(\omega) = (\xi_1(\omega), \ldots, \xi_n(\omega))$ is a realization of random demands; $\xi_{[1,i]}(\omega) = (\xi_1(\omega), \ldots, \xi_i(\omega))$ comprises the demand realizations known at stage i.
- $\mathbf{x}(\omega) = (x_1(\omega), \ldots, x_n(\omega))$ comprises any feasible vector of decisions to be made in response to the random demands $\xi(\omega)$. Thus, $\mathbf{x}(\cdot)$ is a mapping from the set of elementary outcomes Ω to $\mathbf{R}^n$.
- This mapping must nevertheless satisfy the last set of constraints called the *nonanticipativity constraints*, which require that any decision at any stage i can depend only on the information available up to this stage. In other words, $\mathbf{x}(\cdot)$ is not allowed to fit decisions in earlier stages to the information content out from $\xi(\omega)$ which becomes available in later stages. Hence, for any two random outcomes ω^1 and ω^2 such that the demand vectors $\xi(\omega^1)$ and $\xi(\omega^2)$ share the same history $\xi_{[1,i]}$ up to the stage i, the resulting vectors $\mathbf{x}(\omega^1)$ and $\mathbf{x}(\omega^2)$ have to coincide in the stages $\mathbf{x}_{[1,i]}$.

Remember that the expectation of the optimal objective value in the nested formulation (20) as well as in the dynamic programming formulation (23) delivers the expected overall performance $E|n|(s)$. Hence, the same does the expected optimal objective value in the above formulation (28). How to express it? As one can see, we just have to let the demand $\xi_1(\omega)$ become random; the decision x_1 becomes not a deterministic but, respectively, a random variable $x_1(\omega)$. Then, the optimal objective value in (28) will return the expected overall performance $E|n|(s)$.

We can also notice here that this formulation gives us an easier way to prove the Proposition 1, either.

Further, we can eliminate the nonanticipativity conditions from (28) if we require explicitly that the components $x_i(\omega)$ of a decision vector $\mathbf{x}(\omega)$ can depend on the random data observed up to stage i, but not on the future observations, i.e., we require: $x_i(\omega) = x_i(\xi_{[1,i]}(\omega))$. So a decision $x_i = x_i(\xi_{[1,i]})$ is viewed as a function of $(\xi_1, \ldots, \xi_n)$, and maximization is performed over appropriate functional spaces [43, pp. 19, 36], [41, pp. 93–94]. Rearranging the terms in the constraints and eliminating redundant constraints, we can finally transform (28) to:

$$\begin{aligned} \max_{\mathbf{x}(\cdot)} \; E\left[\frac{x_1}{\xi_1} + \frac{x_2}{\xi_2} + \ldots + \frac{x_n}{\xi_n}\right] & \\ x_1 + x_2 + \ldots + x_n \le s & \\ 0 \le x_i(\xi_1, \ldots, \xi_i) \le \xi_i, \quad i = 1, \ldots, n & \end{aligned} \tag{29}$$

We can now formulate and prove the following

Proposition 2 *(Advantage of mobile solution)*

$$\forall s \ge 0 \quad : \quad E|n_1|\ldots|n_K|(s) \le E\mathcal{M}|n_1|\ldots|n_K|(s)$$

Proof:

Indeed, let us refer to problem (29). Its optimal objective value equals $E|n|(s)$. Thus, in order to compute $E|n_1|\ldots|n_K|(s)$ we have to consider the optimal objective values of K instances of (20). These K instances can be unified into one aggregate problem as follows. Let $\mathcal{I} = \{I_1, \ldots, I_K\}$ be the partition of the set $\{1, \ldots, N\}$ into customer groups, with $n_1 = |I_1|, \ldots, n_K = |I_K|$. Then the aggregate problem can be written as:

$$\max_{\mathbf{x}(\cdot)} \; E\left[\frac{x_1}{\xi_1} \quad + \quad \frac{x_2}{\xi_2} \quad + \quad \ldots \quad + \quad \frac{x_n}{\xi_N}\right]$$
$$\sum\nolimits_{i\in\mathcal{I}_1} x_i \;\leq\; s_1$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$\sum\nolimits_{i\in\mathcal{I}_K} x_i \;\leq\; s_K$$
$$0 \;\leq\; x_i(\xi_j : j \leq i, j \in \mathcal{I}_k) \;\leq\; \xi_i, \quad i \in \mathcal{I}_k, \quad k = 1, \ldots, K$$

From the other side, computing $E\mathcal{M}|n_1|\ldots|n_K|(s)$, we need to consider the optimal objective value of a single instance of (29) with $n = N$. As one can see, its constraints is the relaxation of those of the above aggregate problem due to the following:

- Let us consider any mapping $x_i(\xi_{j_1}, \xi_{j_2}, \ldots, \xi_i)$ feasible (implementable) in the aggregate problem. It corresponds to the non-mobile decision making at some stage in some customer group. We see that it can be implemented in (29), too — by means of the mapping $x_i(\xi_1, \ldots, \xi_{j_1}, \ldots, \xi_{j_2}, \ldots, \xi_i)$, which incorporates the same arguments as its non-mobile counterpart, but also, possibly, some others, since we have a longer history $\xi_{[1,i]}$ behind us upon reaching the customer i under mobile decision making.
- Obviously, the first constraint in (29) weakens the first K constraints of the aggregate problem.

Hence, in the mobile case we compute the optimal objective value over a broader (at least not narrower) set of feasible mappings, what proves the proposition. □

To prove this and to quantify this advantage of mobile decision making for any system configuration stays our objective. As an open question stays for us also solving our multistage decision problems modeled in this section.

4 Economic effects: Bernoulli demand distribution

With the above objective, let us try at first to approach distribution systems of lowest possible dimensions while assuming customer demands to follow a particularly simple probability distribution.

Let us assume the customer demands to follow one and the same probability distribution, which we choose to be the Bernoulli one, i.e., a discrete distribution with the state space $\{0, 1\}$, whereby a customer has the demand 1 with the probability $0 < p < 1$. In other words, the probability mass function of any individual customer demand ξ_i looks as:

x	0	1
$Pr\{\xi_i = x\}$	$1-p$	p

4.1 Configuration C|1|

Let us consider at first the single-customer case C|1| introduced in section 3.1, see page 13. Upon arriving to the customer and learning his demand ξ_1 the representative faces the following decision problem:

$$\begin{aligned} Z_1(\xi_1, s) &= \max \frac{x_1}{\xi_1} \\ x_1 &\leq s \\ x_1 &\leq \xi_1 \\ x_1 &\geq 0 \end{aligned}$$

We know how to solve this deterministic linear program: the greatest feasible value of the variable maximizes the objective function. Then, the two possible values of ξ_1 induce the following optimal objective values:

ξ_1 :	0	1
Z_1 :	1	$\min\{s, 1\}$

Then the expectation $E\big[Z_1(\xi_1, s)\big]$ expresses in the case $0 \leq s \leq 1$ as:

$$E[Z_1] = 1 \cdot (1-p) + s \cdot p = 1 - p + sp = 1 - p(1-s).$$

Obviously, in the case $s > 1$ we have $E[Z_1] = 1$.

So, the expected overall performance expresses as:

$$E|1|(s) = E\big[Z_1(\xi_1, s)\big] = \begin{cases} 1 - p \cdot (1-s), & 0 \leq s \leq 1 \\ 1, & s > 1 \end{cases} \tag{30}$$

4.2 Configurations C|1|1| and C$\mathcal{M}$|1|1|

We introduced and discussed these configurations with 2 customers and 2 representatives in section 3.2, see page 14. Let us see what optimal decisions are to be there, and evaluate the average overall performance on the long run in both configurations.

4.3 Scenario C|1|1|

As already shown in section 3.2.1, the 1st and the 2nd representatives face in configuration C|**1**|**1**| once a week the problems (6.1) and (6.2), respectively. Hence, each of them behaves as the single representative in configuration C|**1**|. If the individual stock of each representative amounts to $a = s/2$, then, according to formula (26), we have the expected overall performance

$$E|\mathbf{1}|\mathbf{1}|(s) \;=\; E|\mathbf{1}|(a) \;+\; E|\mathbf{1}|(a) \;=\; 2 \cdot E|\mathbf{1}|(a).$$

Referring to formula (30) for $E|\mathbf{1}|(s)$, we obtain:

$$E|\mathbf{1}|\mathbf{1}|(s) \quad = \quad \begin{cases} 2 - p \cdot (2 - s), & 0 \le s \le 2 \\ 2, & s > 2 \end{cases} \tag{31}$$

4.4 Scenario C$\mathcal{M}$|1|1|

In the mobile scenario C$\mathcal{M}$|**1**|**1**| the representatives achieve the performance of a single "virtual" representative, what we stated by the formula (27). As already shown in the section 3, the 1st representative, who arrives (without loss of generality) to his customer first, faces the optimization problem (8), which we rewrite here in a slightly modified form:

$$\begin{aligned} Z_2(\xi_1, s) \;&=\; \max \frac{x_1}{\xi_1} \;+\; Q_1(x_1) \\ & x_1 \;\le\; s \\ & x_1 \;\le\; \xi_1 \\ & x_1 \;\ge\; 0 \\ & Q_1(x_1) \;=\; E\big[Q_1(x_1, \xi_2)\big] \\ & Q_1(x_1, \xi_2) \;=\; \max_{x_2} \left\{ \frac{x_2}{\xi_2} \;\middle|\; x_2 \le s - x_1,\; x_2 \le \xi_2,\; x_2 \ge 0 \right\} \end{aligned} \tag{32}$$

The only decision variable here is x_1: the allocation to the 1st customer is being made under uncertainty concerning the 2nd customer's demand.

How are we going to treat the function $Q_1(x_1)$? As one may see, this is the expectation of 2nd representative's performance under condition that there are $s - x_1$ units of the good left at his disposal, and, hence,

$$Q_1(x_1) = E\big[Q_1(x_1, \xi_2)\big] = E\big[Z_1(\xi_1, s - x_1)\big] = E|\mathbf{1}|(s - x_1).$$

Referring to formula (30) for $E|\mathbf{1}|(s)$ and substituting $s - x_1$ for s there, we obtain the following expression of $Q_1(x_1)$:

$$Q_1(x_1) = \begin{cases} 1 - p(1 - s + x_1), & 0 \le s - x_1 \le 1 \\ 1, & s - x_1 > 1 \end{cases} = \begin{cases} 1 - p + p(s - x_1), & s - 1 \le x_1 \le s \\ 1, & x_1 < s - 1 \end{cases} \tag{33}$$

Now the problem (32) can be rewritten as

$$Z_2(\xi_1, s) \;=\; \max \left\{ \frac{x_1}{\xi_1} + Q_1(x_1) \;\middle|\; x_1 \in [0, \min\{\xi_1, s\}] \right\}$$

where, as always, $\xi_1 = 0$ implies $x_1 = 0$ and the corresponding fraction equals 1 per definition.

How can we obtain the expected overall performance $E\mathcal{M}|1|1|(s)$? As stated by Proposition 1 in section 3.4, the expected overall performance in this configuration is given by the expected value of $Z_2(\xi_1, s)$. Hence, we can compute $E\mathcal{M}|1|1|(s)$ in a similar way as we computed $E|1|(s)$: since the random parameter ξ_1 has a finite state space $\{0, 1\}$, we may determine the optimal objective value $Z_2(\xi_1, s)$ for each value of the random parameter, and then express the expectation. Let us compute the optimal objective value $Z_2(\xi_1, s)$ for each possible realization of ξ_1:

(a) $\xi_1 = 0 \;\Rightarrow\;$ according to the constraints in (32), we have an only feasible value of the variable: $x_1 = 0$. Hence, the optimal objective value expresses as:

$$Z_2(0, s) = 1 + Q_1(0) = 1 + \begin{cases} 1 - p(1 - s), & 0 \le s \le 1 \\ 1, & s > 1 \end{cases} = \begin{cases} 2 - p(1 - s), & 0 \le s \le 1 \\ 2, & s > 1 \end{cases}$$

(b) $\xi_1 = 1 \;\Rightarrow\;$ we have to maximize the objective function on the closed interval:

$$\frac{x_1}{1} + Q_1(x_1) \quad \longrightarrow \quad \max_{[0, \min\{1, s\}]}$$

As one may see, the objective function is continuous (piecewise linear) and monotonically increasing in x_1. Hence, the greatest possible value of the variable delivers the optimum to the objective function:

$$x_1^* = \min\{1, s\} = \begin{cases} s, & s \le 1 \\ 1, & s > 1 \end{cases}$$

$$Z_2(1, s) = x_1^* + Q_1(x_1^*) = \begin{cases} s + 1 - p, & s \le 1 \\ 1 + \left\{ \begin{array}{ll} 1 - p + p(s - 1), & 0 \le s - 1 \le 1 \\ 1, & s - 1 > 1 \end{array} \right\}, & s > 1 \end{cases} =$$

$$= \begin{cases} s + 1 - p, & s \le 1 \\ 2 - p(2 - s), & 1 < s \le 2 \\ 2, & s > 2 \end{cases}$$

So, the optimal objective values are known under all possible realizations of the random parameter ξ_1. Hence, the distribution of $Z_2(\xi_1, s)$ is known to us, and we can express its expectation $E\big[Z_2(\xi_1, s)\big]$:

$$E\big[Z_2(\xi_1,s)\big] \quad = \quad Pr\{\xi_1=0\}\cdot Z_2(0,s) \;+\; Pr\{\xi_1=1\}\cdot Z_2(1,s).$$

We split the evaluation of this expression into three parts:

(1) $0 \le s \le 1$:

$$\begin{aligned} E[Z_2] &= (1-p)\cdot(2-p(1-s)) + p\cdot(s+1-p) = \\ &= 2-2p-p+p^2+ps-p^2s+ps+p-p^2 = 2(1-p)+ps(2-p). \end{aligned}$$

(2) $1 < s \le 2$:

$$E[Z_2] = (1-p)\cdot 2 + p\cdot(2-p(2-s)) = 2-2p+2p-2p^2+p^2s = 2-p^2(2-s).$$

(3) $s > 2$:

$$E[Z_2] = (1-p)\cdot 2 + p\cdot 2 = 2.$$

Then, we can summarize: the average overall performance of the two representatives in the mobile scenario $C\mathcal{M}|1|1|$ with a stock s replenished weekly, equals on the long run that of one single representative in scenario C$|2|$, which both express as:

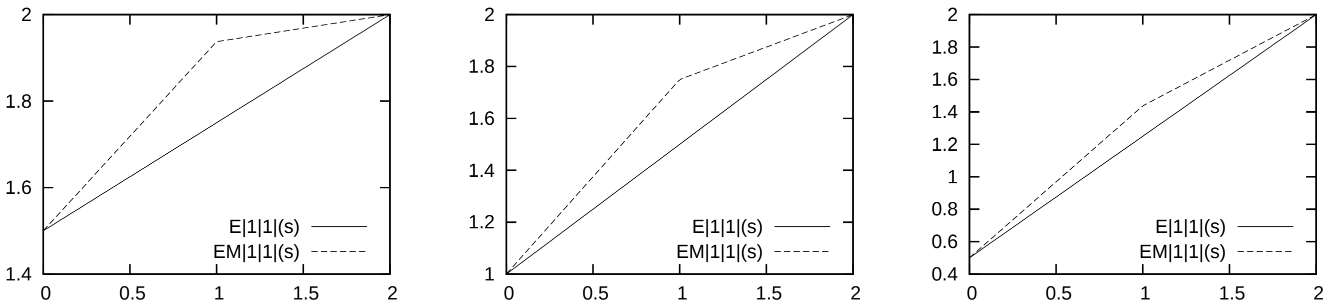
$$E|2|(s) \;=\; E\mathcal{M}|1|1|(s) \;=\; \begin{cases} 2(1-p)+ps(2-p), & 0 \le s \le 1 \\ 2-p^2(2-s), & 1 < s \le 2 \\ 2, & s > 2 \end{cases} \tag{34}$$

4.5 Comparing C$|1|1|$ and C$\mathcal{M}|1|1|$

We compare now here the overall performance of the representatives in scenarios C$|1|1|$ and C$\mathcal{M}|1|1|$ according to the formulas (31) and (34) obtained above. Figure 9 displays the two functions $E|1|1|(s)$ and $E\mathcal{M}|1|1|(s)$ for some particular values of p: we chose, for instance, $p = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$.

As Proposition 2 states, the advantage of mobile configuration holds in general. Still, one can verify the inequality $E|1|1|(s) \le E\mathcal{M}|1|1|(s)$ explicitly by comparing the expressions of these two functions. Such verification can show how interrelationships between parameters make the functions differ from each other, and what parameter values would make them equal. We have presented this verification in Appendix A.

The advantage of C$\mathcal{M}|1|1|$ can also be expressed in per cent relative to C$|1|1|$, as $[E\mathcal{M}|1|1| - E|1|1|]/E|1|1|$. We display it graphically for the chosen above demand probabilities $p = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ in Figure 10.



$p = \frac{1}{4}$ $p = \frac{1}{2}$ $p = \frac{3}{4}$

Figure 9: Comparison of mobile and non-mobile performances

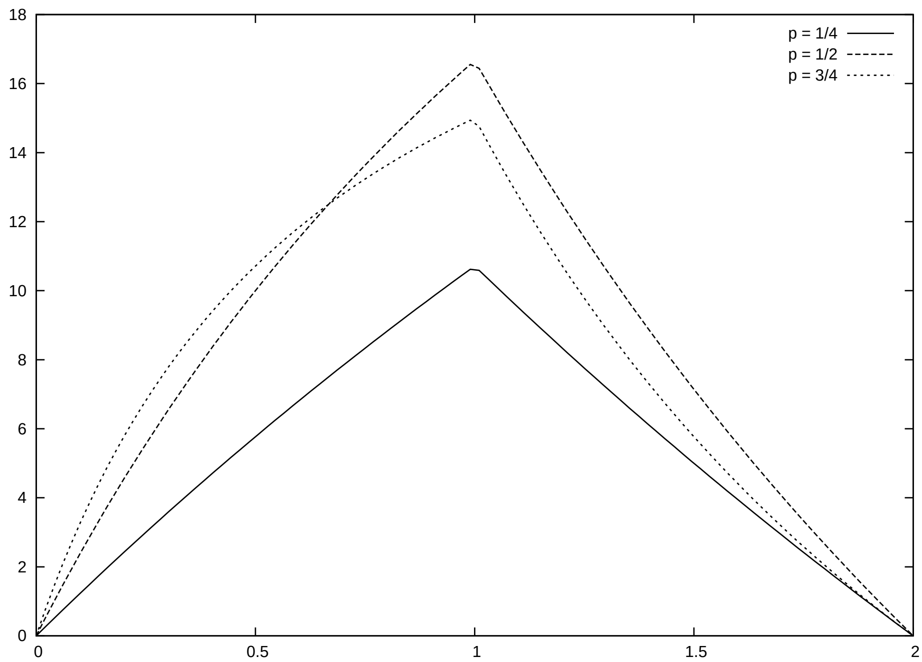


Figure 10: Relative advantage of mobile solution $\mathrm{C}\mathcal{M}|1|1|$ at different demand probabilities (in %)

4.6 Servicing 3 customers

We are going to quantify now the effects of the mobile solution in a distribution system that incorporates 3 customers. Thereby we can distinguish between the cases where the whole job is carried out by either 2 or 3 representatives.

4.6.1 Servicing 3 customers by 3 representatives

We compare here configurations $\mathrm{C}|1|^3$ and $\mathrm{C}\mathcal{M}|1|^3$, which we introduced and discussed in section 3.3.1, see page 19.

Let us refer at first to the non-mobile configuration $\mathrm{C}|1|^3$. Three identical representatives, working independently of each other, receive equal quotas from the common stock s to serve their customers: $s = 3a$. Hence, the expected overall performance in this case can be found as

$$E|1|^3(s) \;=\; 3 \cdot E|1|(s/3).$$

Referring to formula (30) for $E|1|(s)$, we obtain:

$$E|1|^3(s) \;=\; \begin{cases} 3 - p\cdot(3-s), & 0 \le s \le 3 \\ 3, & s > 3 \end{cases} \tag{35}$$

Let us now refer to the mobile configuration $\mathrm{C}\mathcal{M}|1|^3$. The decision making of the 1st representative is represented in this case by the 3-stage stochastic program (16), which we rewrite here in a slightly modified form:

$$\begin{aligned}
Z_3(\xi_1, s) \;&=\; \max \frac{x_1}{\xi_1} \;+\; Q_2(x_1) \\
&\quad x_1 \;\le\; s \\
&\quad x_1 \;\le\; \xi_1 \\
&\quad x_1 \;\ge\; 0 \\
&\quad Q_2(x_1) = E\big[Q_2(x_1,\xi_2)\big] \\
&\quad Q_2(x_1,\xi_2) \;=\; \max_{x_2}\left\{ \frac{x_2}{\xi_2} + E\left[\max_{x_3}\frac{x_3}{\xi_3}\right] \;\middle|\; \begin{array}{ll} x_2 \;\le\; s - x_1, & x_2 \;\le\; \xi_2,\; x_2 \;\ge\; 0 \\ x_3 \;\le\; s - x_1 - x_2, & x_3 \;\le\; \xi_3,\; x_3 \;\ge\; 0 \end{array}\right\}
\end{aligned} \tag{36}$$

As we already know, function $Q_2(x_1)$ proves to be found as $E\mathcal{M}|1|1|(s - x_1)$, since it expresses the expected performance of the 2nd and 3rd representatives with the stock $s - x_1$ left after their 1st colleague's decision. Then, referring to formula (34) for $E\mathcal{M}|1|1|(s)$ and substituting $s - x_1$ for s there, we obtain the following expression of $Q_2(x_1)$:

$$\begin{aligned}
Q_2(x_1) \;&=\; \begin{cases} 2(1-p) + p(s-x_1)(2-p), & 0 \le s - x_1 \le 1 \\ 2 - p^2(2 - (s - x_1)), & 1 < s - x_1 \le 2 \\ 2, & s - x_1 > 2 \end{cases} \\
&=\; \begin{cases} 2(1-p) + p(s-x_1)(2-p), & s - 1 \le x_1 \le s \\ 2 - p^2(2 - (s - x_1)), & s - 2 \le x_1 < s - 1 \\ 2, & x_1 < s - 2 \end{cases}
\end{aligned} \tag{37}$$

Now the problem (36) can be rewritten as

$$Z_3(\xi_1, s) \;=\; \max\left\{ \frac{x_1}{\xi_1} + Q_2(x_1) \;\middle|\; x_1 \in [0, \min\{\xi_1, s\}] \right\}$$

where, as always, $\xi_1 = 0$ implies $x_1 = 0$ and the corresponding fraction equals 1 per definition.

How to express the average overall performance in this scenario on the long run? As we already know, the expected overall performance in this configuration is given by the expectation of the optimal objective value $Z_3(\xi_1, s)$. Again, we can express $E\mathcal{M}|1|^3(s)$ by computing $Z_3(\xi_1, s)$ for each possible realization of the random parameter ξ_1. Let us do this.

One can easily do this since the objective function in (4.6.1) is continuous, piecewise linear and monotonically increasing function (these properties can be verified directly). Hence, the greatest possible value of the argument will deliver an optimum to the objective function.

(a) $\xi_1 = 0 \;\Rightarrow\;$ we have the only feasible value of the variable: $x_1 = 0$. Hence, the optimal objective value expresses as:

$$\begin{aligned} Z_3(0) &= 1 + Q_2(0) = \\ &= 1 + \begin{cases} 2(1-p) + ps(2-p), & 0 \le s \le 1 \\ 2 - p^2(2-s), & 1 < s \le 2 \\ 2, & s > 2 \end{cases} = \begin{cases} 3 - 2p + ps(2-p), & 0 \le s \le 1 \\ 3 - p^2(2-s), & 1 < s \le 2 \\ 3, & s > 2 \end{cases} \end{aligned}$$

(b) $\xi_1 = 1:$ Since the greatest possible value of the variable is optimal, we have:

$$x_1^* = \min\{1, s\} = \begin{cases} s, & s \le 1 \\ 1, & s > 1 \end{cases}$$

$$Z_3(1) = x_1^* + Q_2(x_1^*) = \begin{cases} s + 2(1-p), & s \le 1 \\ 1 + \begin{cases} 2(1-p) + p(s-1)(2-p), & 0 \le s-1 \le 1 \\ 2 - p^2(2-(s-1)), & 1 < s-1 \le 2 \\ 2, & s-1 > 2 \end{cases}, & s > 1 \end{cases}$$

$$= \begin{cases} s + 2(1-p), & s \le 1 \\ 3 - 2p + p(s-1)(2-p), & 1 < s \le 2 \\ 3 - p^2(3-s), & 2 < s \le 3 \\ 3, & s > 3 \end{cases}$$

So, the optimal objective values are known under all possible realizations of the random parameter ξ_1. Hence, the distribution of $Z_3(\xi_1, s)$ is known to us, and we can express its expectation $E\left[Z_3(\xi_1, s)\right]$:

$$E\mathcal{M}|1|^3(s) \;=\; E\big[Z_3(\xi_1, 3)\big] \;=\; Pr\{\xi_1 = 0\} \cdot Z_3(0) \;+\; Pr\{\xi_1 = 1\} \cdot Z_3(1).$$

We split the evaluation of this expression into four parts:

(1) $0 \le s \le 1:$

$$E[Z_3] = (1-p) \cdot (3 - 2p + ps(2-p)) + p \cdot (s + 2(1-p)) =$$

$$= 3 - 3p + 3ps - 3p^2s + p^3s = 3 - 3p + ps(p^2 - 2p + 1 - p + 2) =$$
$$= 3 - 3p + ps(2 - p) + ps(1 - p)^2.$$

(2) $1 < s \leq 2$:

$$E[Z_3] = (1 - p) \cdot (3 - p^2(2 - s)) + p \cdot (3 - 2p + p(s - 1)(2 - p)) =$$
$$= 3 - 6p^2 + 3p^2s + 3p^3 - 2p^3s = 3 - 3p^2(2 - s) - p^3(2s - 3).$$

(3) $2 < s \leq 3$:

$$E[Z_3] = (1 - p) \cdot 3 + p \cdot (3 - p^2(3 - s)) = 3 - 3p^3 + p^3s = 3 - p^3(3 - s).$$

(4) $s > 3$:

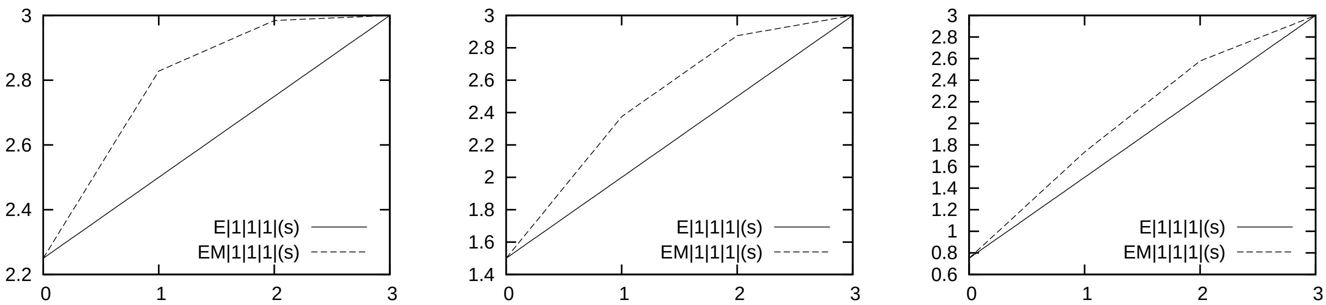
$$E[Z_3] = (1 - p) \cdot 3 + p \cdot 3 = 3.$$

Then, we can summarize: the average overall performance of the three representatives in the mobile scenario $C\mathcal{M}|1|^3$, provided with a fixed stock s weekly, expresses on the long run as

$$E\mathcal{M}|1|^3(s) \;=\; E|3|(s) \;=\; \begin{cases} 3 - 3p + ps(2 - p) + ps(1 - p)^2, & 0 \leq s \leq 1 \\ 3 - 3p^2(2 - s) - p^3(2s - 3), & 1 < s \leq 2 \\ 3 - p^3(3 - s), & 2 < s \leq 3 \\ 3, & s > 3 \end{cases} \tag{38}$$

4.6.2 Comparing C$|1|^3$ and C$\mathcal{M}|1|^3$

We compare now here the overall performance of the representatives in scenarios $C|1|^3$ and $C\mathcal{M}|1|^3$ according to the formulas (35) and (38) obtained above. Figure 11 displays two functions $E|1|^3(s)$ and $E\mathcal{M}|1|^3(s)$ for $p = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$.



$p = \dfrac{1}{4}$ $\qquad$ $p = \dfrac{1}{2}$ $\qquad$ $p = \dfrac{3}{4}$

Figure 11: Comparison of mobile and non-mobile performances of 3 representatives

Again, as Proposition 2 states, the advantage of mobile configuration holds in general: i.e., in our case it stays true for any probability p. Still, one can verify the inequality $E|1|^3(s) \leq E\mathcal{M}|1|^3(s)$ explicitly by comparing the expressions of these two functions. Such verification can show how interrelationships between parameters make the functions differ from each other, and what parameter values would make them equal. We have presented this verification in Appendix A.

The advantage of C$\mathcal{M}|1|^3$ can also be expressed in per cent relative to C$|1|^3$, as $[E\mathcal{M}|1|^3 - E|1|^3]/E|1|^3$. We display it graphically for the chosen above demand probabilities $p = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ in Figure 12.

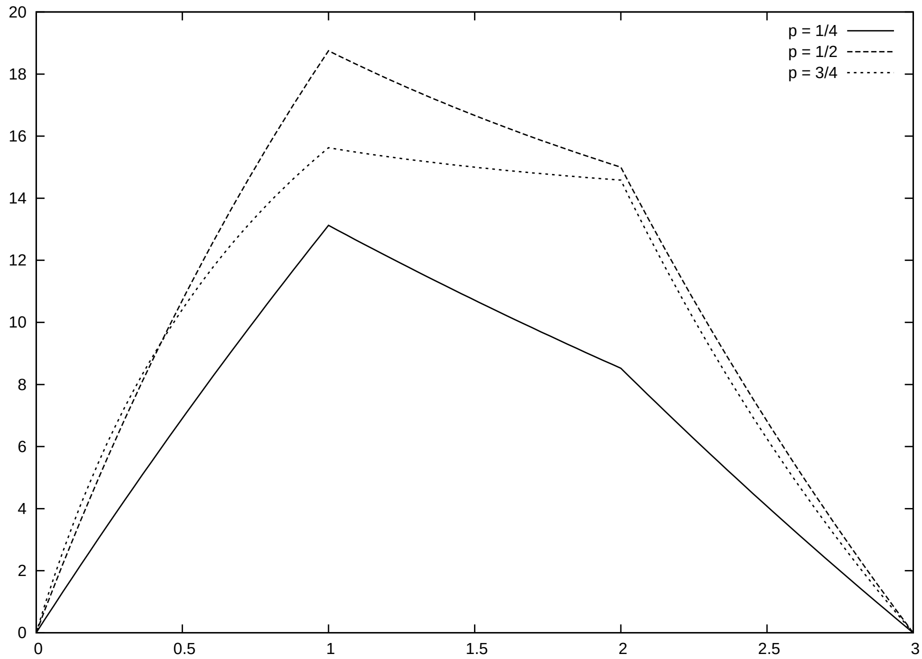


Figure 12: Relative advantage of mobile solution C$\mathcal{M}|1|^3$ comparing to C$|1|^3$ at different demand probabilities (in %)

4.6.3 Servicing 3 customers by 2 representatives

We consider here configurations C$|2|1|$ and C$\mathcal{M}|2|1|$introduced and discussed in section 3.3.2, see page 22. We sketch them here once again in Figure 13.

Decision making in both configurations is already known to us: in C$|2|1|$ it is carried out as in C$|1|1|$ and C$|1|$, while in C$\mathcal{M}|2|1|$ it is equivalent to the case C$|3|$, and hence, to C$\mathcal{M}|1|^3$ — which we have analyzed in the previous section.

The overall performance express in the non-mobile case C$|2|1|$ as:

$$E|2|1|(s) \;=\; E|2|\left(\frac{2}{3}s\right) + E|1|\left(\frac{1}{3}s\right) \;=$$

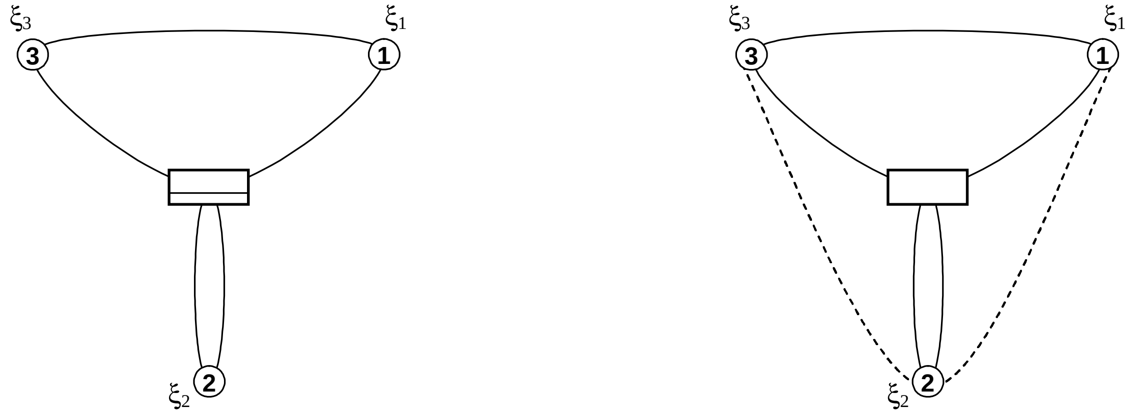


Figure 13: Configurations C|2|1| and C$\mathcal{M}$|2|1|

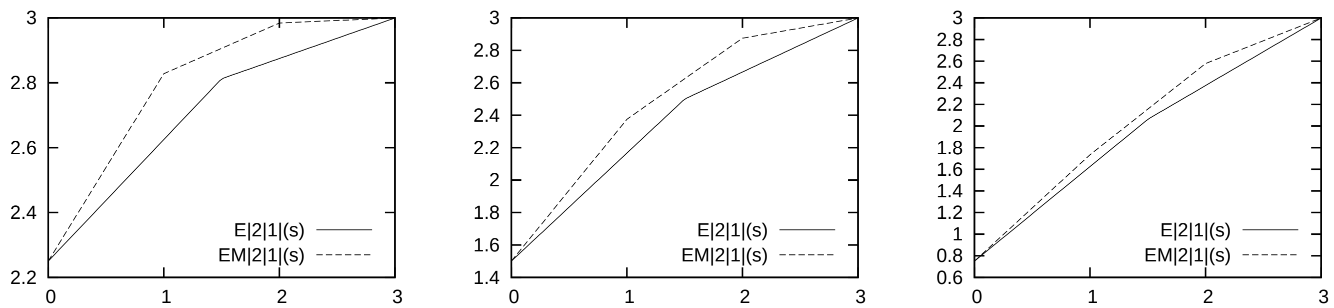
$$= \begin{cases} 2(1-p) + p \cdot \frac{2}{3}s(2-p), & 0 \le \frac{2}{3}s \le 1 \\ 2 - p^2(2 - \frac{2}{3}s), & 1 < \frac{2}{3}s \le 2 \\ 2, & \frac{2}{3}s > 2 \end{cases} + \begin{cases} 1 - p(1 - \frac{1}{3}s), & 0 \le \frac{1}{3}s \le 1 \\ 1, & \frac{1}{3}s > 1 \end{cases} =$$

$$= \begin{cases} 2 - 2p + \frac{2}{3}ps(2-p), & 0 \le s \le \frac{3}{2} \\ 2 - p^2(2 - \frac{2}{3}s), & \frac{3}{2} < s \le 3 \\ 2, & s > 3 \end{cases} + \begin{cases} 1 - p + \frac{1}{3}ps, & 0 \le s \le 3 \\ 1, & s > 3 \end{cases} =$$

$$= \begin{cases} 3 - 3p + \frac{5}{3}ps - \frac{2}{3}p^2 s, & 0 \le s \le \frac{3}{2} \\ 3 - (2p^2 + p)(1 - \frac{1}{3}s), & \frac{3}{2} < s \le 3 \\ 3, & s > 3 \end{cases} \tag{39}$$

In C$\mathcal{M}$|2|1| the expected overall performance coincides with $E|3|(s)$ given by formula (38).

Figure 14 displays the overall performance on the long run in both scenarios graphically, for some particular demand probabilities $p = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$.



$p = \dfrac{1}{4}$ $\qquad$ $p = \dfrac{1}{2}$ $\qquad$ $p = \dfrac{3}{4}$

Figure 14: Mobile and non-mobile performances of 2 representatives servicing 3 customers

The advantage of mobile configuration holds in general — as stated by Proposition 2.

Figure 15 displays the relative advantage $\big(E\mathcal{M}|2|1| - E|2|1|\big)/E|2|1|$, for the chosen above demand probabilities $p = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$:

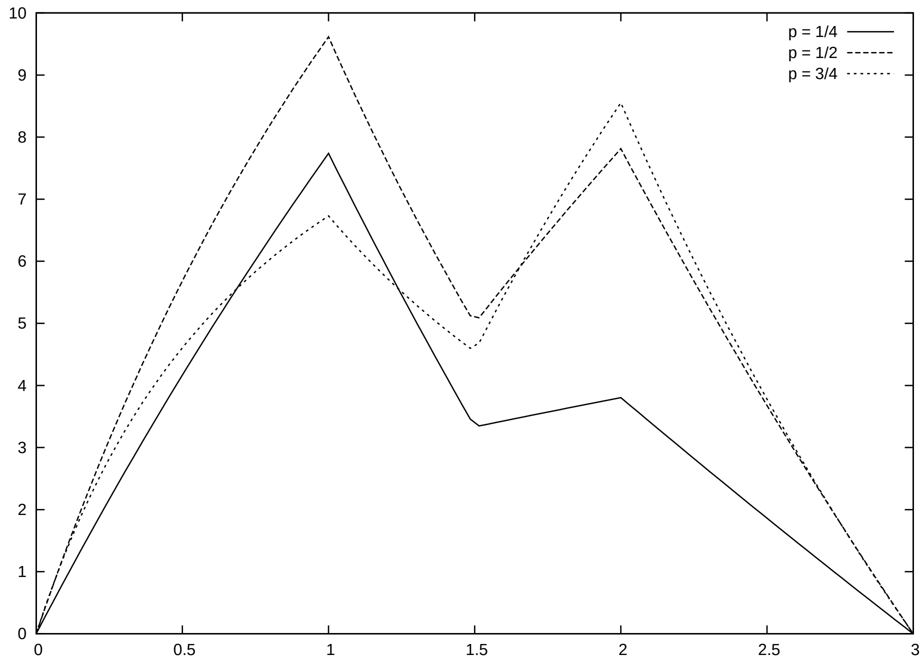


Figure 15: Relative advantage of mobile solution $C\mathcal{M}|2|1|$ comparing to $C|2|1|$ at different demand probabilities (in %)

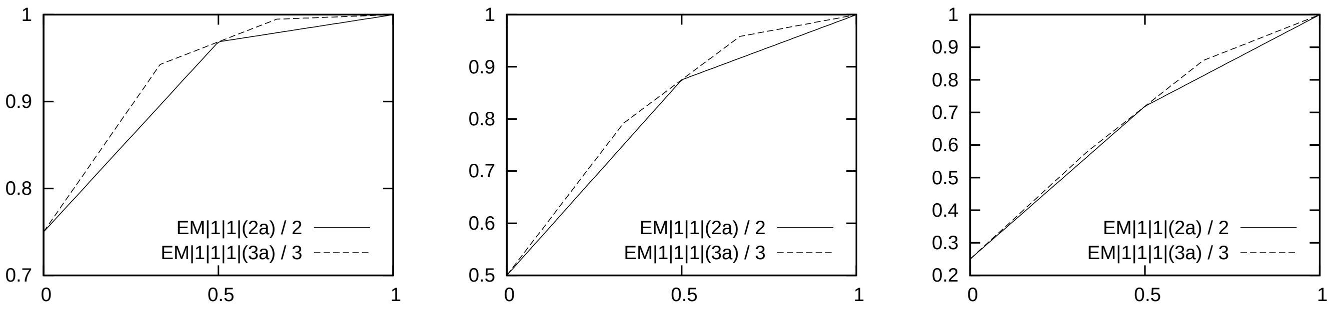
As one may see, this advantage of $C\mathcal{M}|2|1|$ comparing to $C|2|1|$ is essentially lower than the advantage of $C\mathcal{M}|1|^3$ comparing to $C|1|^3$, namely due to the fact that one of the representatives in the non-mobile scenario $C|2|1|$ services two customers alone, what makes him more efficient than 2 independent representatives servicing the same two customers in the non-mobile scenario $C|1|^3$. This makes $C|2|1|$ closer to its mobile counterpart $C\mathcal{M}|2|1|$, and so $C\mathcal{M}|2|1|$ shows a lesser advantage.

4.7 Average mobile performance per customer

After we have analyzed the performance of the representatives in all scenarios from $C|1|1|$ to $C\mathcal{M}|2|1|$, let us raise the following question: having identical customers everywhere in all these scenarios, and setting always the initial weekly stock *per customer* to a, what is the average performance of the representatives per customer then? We are especially interested in comparing the mobile scenarios $C\mathcal{M}|1|1|$ and $C\mathcal{M}|1|^3$ here, since they incorporate different numbers of customers: 2 and 3, respectively. Hence, the total weekly stocks amount in these scenarios to $2a$ and $3a$, respectively. Thus, we are going to compare $E\mathcal{M}|1|1|(2a)/2$ and $E\mathcal{M}|1|^3(3a)/3$, and we display at first these specific performances graphically for the demand probabilities $p = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ in Figure 16.

As we see, an average customer gets in most cases a better service in scenario $C\mathcal{M}|1|^3$ — at least at the chosen demand probabilities.

Let us also display the relative service improvement $\left[\frac{E\mathcal{M}|1|^3(3a)/3}{E\mathcal{M}|1|1|(2a)/2} - 1\right]$ which an average customer gets in $C\mathcal{M}|1|^3$ comparing to $C\mathcal{M}|1|1|$, depending on how much the company reserves for him weekly. We do this for the chosen above demand probabilities $p = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ in Figure 17.



$p = \frac{1}{4}$ $p = \frac{1}{2}$ $p = \frac{3}{4}$

Figure 16: Average performance per customer in mobile scenarios C$\mathcal{M}|1|1|$ and C$\mathcal{M}|1|^3$

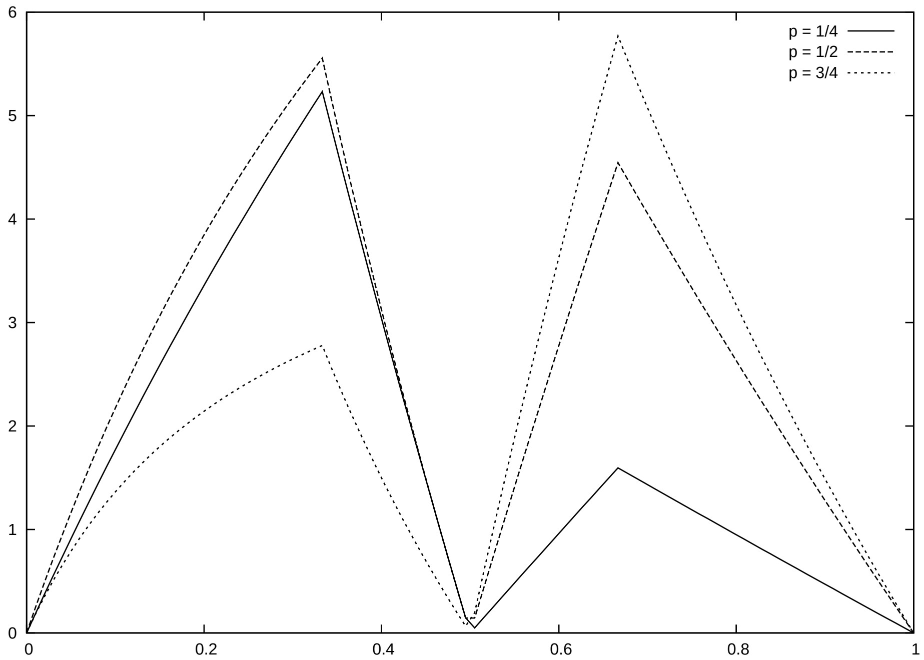


Figure 17: Service improvement for the average customer in C$\mathcal{M}|1|^3$ comparing to C$\mathcal{M}|1|1|$ at different demand probabilities (in %)

So such customer integration into a more global resource pool promises more to an average customer. We might expect such effect in configurations of higher dimensions: if the representatives service their assigned territories independently of each other, i.e. each provides his service sequentially to the customers within a particular cluster, then the consolidation of all resources into a common pool and allocation from this pool by means of mobile technologies globally might lead to a better service without any change of the weekly stock s!

But, we have to discuss here another important question. We are speaking about the average customer (who presumably gets a better service), but who is he? He is rather an imaginary customer. Can we say that the service improves for all customers somehow equally, namely by a percentage like the one given by the Figure 17?

The answer is rather no: as one may guess, servicing in a sequence is order-dependent. So it makes difference for a particular customer to be at the end of the sequence or at its beginning: if we recall how the allocation decisions of the representatives in the configurations analyzed above (e.g. C$\mathcal{M}|1|1|$ and C$\mathcal{M}|1|^3$) are made, we will notice that under Bernoulli demand distribution the decision was always to give the customer as much as possible (remember how it sounded: "the greatest possible value of the argument delivers the optimum to the objective function"). Hence, the customers at early stages will always have in average a better service, and the farther a particular customer from the beginning of the sequence, the worse his servicing is expected to be then, since the probability that the stock has been essentially exhausted at the preceding stages (regardless of the other customers) can be only higher. So we might expect that the earlier stages are more "profitable" then the later ones, and there might be customers in the system for whom the integration in a global mobile servicing means definitely worsening of service even though the overall system performance improves. And so we can formulate our further research questions here:

Q4: How could we provide a more fair resource allocation? By which mechanism?

Q5: Does such "greedy" allocation policy, which were observing in this section, stay always optimal, or not?

As we will see later, the answer to Q5 is no: not always, there are demand distributions, under which the allocations have to be made moderately. Concerning Q4, we might suggest to alternate weekly the sequential ordering (e.g. by a cyclic shift), making the customers appear equally frequent at the beginning as well as at the end of the sequence. But, unfortunately, this solution seems to be impractical, and we have to strive to find more appropriate ones.

5 Economic effects: any finite discrete distribution

Now, after we went in section 4 through a number of configurations, our approach in analyzing effects of mobile technologies seems to be sketched, and we could apply it now to distribution systems of higher dimensions: we would have to continue resolving the function $E|n|(s)$ recursively for greater and greater values of n, until all required $E|n_k|(s)$ and, finally, $E|N|(s)$ become known.

Let us now consider customer demands that follow some arbitrary discrete distribution with finite support. I.e., each demand $\xi_i,\ i = 1, \ldots, N$, is assumed now to have the following probability mass function (pmf):

$$\begin{array}{c||c|c|c} x & d_1 & \ldots & d_m \\ \hline Pr\{\xi_i = x\} & p_1 & \ldots & p_m \end{array} \tag{40}$$

with $0 \leq d_1 < \ldots < d_m$.

Do we have another way — other than recursive computation — to obtain the optimal first-stage decision x_1^*, as well as to express the expected overall performance $E|n|(s)$, for any n?

5.1 Forming a deterministic equivalent

Let us refer to formulation (29) of the multistage stochastic linear program, given in section 3.4 on page 29. We repeat this formulation once again here:

$$
\begin{aligned}
\max_{\mathbf{x}(\cdot)} \; E\left[\frac{x_1}{\xi_1} + \frac{x_2}{\xi_2} + \ldots + \frac{x_n}{\xi_n}\right] & \\
x_1 + x_2 + \ldots + x_n \leq s & \\
0 \leq x_i(\xi_1, \ldots, \xi_i) \leq \xi_i, \quad i = 1, \ldots, n &
\end{aligned}
\tag{41}
$$

Recall that this formulation replaces the sequence of nested objective functions with the single one, while the maximization is performed over a set of measurable mappings $x_i(\xi_1, \ldots, \xi_i)$. Its optimal objective value expresses the expected overall performance $E|n|(s)$ prior to learning any customer demand. Consequently, having in the first stage the demand ξ_1 already deterministic, we obtain with (41) the optimal first-stage decision x_1^*.

Obviously, in case of discrete finite demand distributions, we can express the objective function in (41) as a deterministic linear function, as well as represent the entire problem as one large deterministic linear program.

To do this, let us represent all possible realizations of the random data in our problem in the form of a *stochastic decision tree* [41, p. 29], [24, p. 119], [8, pp. 129–130]. Figure 18 gives an example of it assuming $n = 4$ and $\{1, 2\}$ to be the state space of the random demands $\xi_1, \ldots, \xi_4$.

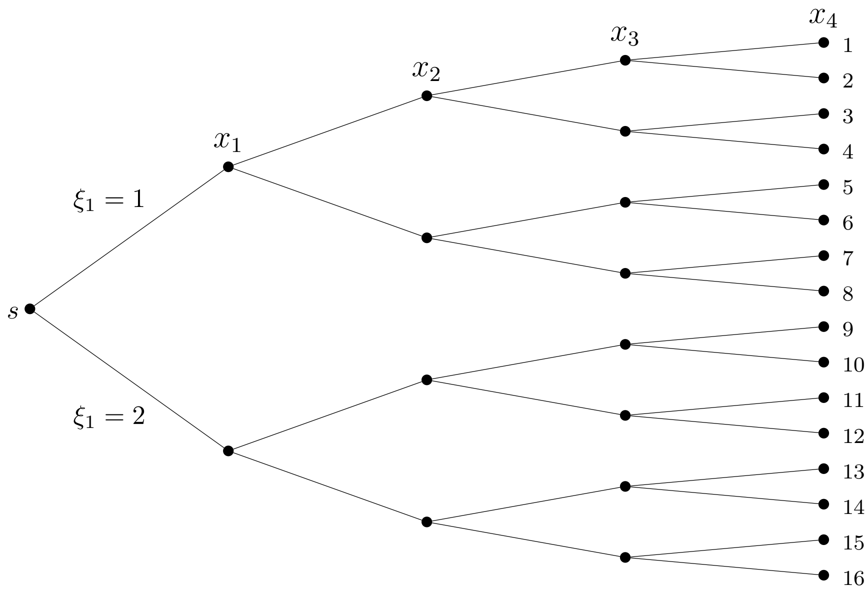


Figure 18: Example of a stochastic decision tree

Let us review this example. In the 1st stage the random data ξ_1 has two possible realizations. We respond to the random demand in this stage by means of the mapping $x_1(\xi_1)$. In the 2nd stage, the random data (ξ_1, ξ_2) has four possible realizations, and we must define response to the corresponding four scenarios by means of the mapping $x_2(\xi_1, \xi_2)$. Then, in the 3rd stage, the random data (ξ_1, ξ_2, ξ_3) is represented by 8 possible scenarios, to which we respond by means of $x_3(\xi_1, \xi_2, \xi_3)$ — having to define 8 values here. On the last stage the number of possible scenarios reaches 16, and the mapping $x_4(\xi_1, \ldots, \xi_4)$ must define 16 values, respectively. Thus, we can associate mapping $x_1(\xi_1)$ with its two values x_1^1, x_1^2, then mapping $x_2(\xi_1, \xi_2)$ with its four values $x_2^{11}, x_2^{12}, x_2^{21}, x_2^{22}$, and so on. Since our (yet unknown) mappings $x_1(\cdot), \ldots, x_4(\cdot)$ are fully defined by their values, we can treat the latter as just variables to which we should assign values optimally. Then, in the case presented by Figure 18, we will have $2 + 4 + 8 + 16 = 30$ variables in total.

Thus, having a finite number of scenarios in each stage to count on, and assigning a decision variable in each stage to each possible scenario, we can rewrite (41) as the following program:

$$
\begin{aligned}
& \max \sum_{t=1}^{M} P_t \cdot \left[\frac{x_1^t}{\xi_1^t} + \frac{x_2^t}{\xi_2^t} + \ldots + \frac{x_n^t}{\xi_n^t} \right] \\
& x_1^t + x_2^t + \ldots + x_n^t \le s, \quad t = 1, \ldots, M \\
& 0 \le x_i^t \le \xi_i^t, \quad i = 1, \ldots, n, \quad t = 1, \ldots, M
\end{aligned}
\tag{42}
$$

where M is the total number of scenarios. Since each scenario corresponds to a path from the root to one of the leaves in the stochastic decision tree, the probability P_t of realization of scenario t is given by

$$P_t = \prod_{i=1}^{n} p_i^t,$$

where p_i^t is the probability for the demand in stage i to follow the path t. I.e., P_t is the product of the probabilities associated with the edges constituting the path t. Respectively, ξ_i^t and x_i^t are the demand realizations and decisions associated with these edges. Formulation (42) is also called *deterministic equivalent* program.

Apparently, a term $P_t \cdot \frac{x_i^t}{\xi_i^t}$ will have to be included in the summation in (42) as many times, as many paths in the stochastic decision tree are traversing through the corresponding edge in the i-th stage. Let us denote the number of edges in stage i of the stochastic decision tree by m_i, and associate decision variables $x_i^j,\ j = 1, \ldots, m_i$ with these edges. Let us also denote by ξ_i^j and p_i^j the demand realizations and the probabilities corresponding to these edges, respectively, and by P_i^j — the probability of traversing through the corresponding edge while starting at the root. Obviously, the latter is expressed as the product of the probabilities of all preceding edges and of this edge p_i^j. Then, we can rewrite (42) also in the following form:

$$
\begin{aligned}
& \max \sum_{j=1}^{m_1} P_1^j \frac{x_1^j}{\xi_1^j} + \ldots + \sum_{j=1}^{m_n} P_n^j \frac{x_n^j}{\xi_n^j} \\
& x_1^{j_1} + \ldots + x_n^{j_n} \le s, \quad j_i = 1, \ldots, m_i, \quad i = 1, \ldots, n \\
& 0 \le x_i^j \le \xi_i^j, \quad i = 1, \ldots, n, \quad j = 1, \ldots, m_i
\end{aligned}
\tag{43}
$$

In this formulation the variables appear in the objective exactly once, since we have grouped together their multiple appearances in the objective of (42). The constraints remain in fact the same: we only have adjusted them to the notation properly.

Let us now estimate the size of the above linear program, having n stages in our sequential decision making and provided with m possible realizations of each customer demand. We have, obviously, in total

$$m_1 + \ldots + m_n$$

decision variables with $m_1 = m$, $m_2 = m^2$, $\ldots$, $m_n = M = m^n$. Hence, the number of the variables equals the sum of geometric series:

$$m \cdot \frac{m^n - 1}{m - 1}.$$

The number of constraints is easier to determine with the formulation (42). Namely, the first set of inequalities provides us with $M = m^n$ constraints, while the second set — with one more constraint and a nonnegativity for each of the variables.

Let us consider, for example, the case with $n = 4$ stages and $m = 2$ possible demand realizations. Then, in terms of a standard linear program of the form

$$\begin{aligned} &\max c^T x \\ &\quad Ax \leq b \\ &\quad x \geq 0 \end{aligned}$$

we have to deal with the matrix A of dimension 46×30. For the sake of convenience, let us nevertheless turn back to the original form of our linear program:

$$\begin{aligned} &\max c^T x \\ &\quad Ax \leq b_1 \\ &\quad 0 \leq x \leq b_2 \end{aligned} \tag{44}$$

Then, assuming the state space of the random demands to be $\{1, 2\}$, we obtain the components A, b_1, b_2, c of our linear program as presented in Figure 19.

The case when the state space contains 0 as a possible demand realization is a special one. As we know, a demand realization $\xi_i = 0$ implies the decision $x_i = 0$, and the corresponding fraction in the objective equals 1 by definition. In this case we have to perform some pre-processing of the model (43). Indeed, those variables that correspond to the edges associated with demands $\xi_i^j = 0$ are known in advance to be equal 0; their contribution to the objective function value is also known in advance. Hence, we can exclude these variables from the model, what also reduces by that its size. Namely, having n stages with identically distributed demands, we reduce the number of variables by n.

It is not hard to develop a procedure that creates the components of the model (43) for any given n and any given finite demand distributions, and then pass this model to any convenient LP solver. By that we can compute the expected overall performance prior to knowing any demand realization, as well as the first decision x_1^* upon learning the first demand realization ξ_1. In the same way we obtain the decision x_2^* as the first one in the sequence of $n-1$ customers while having the stock reduced to $s - x_1^*$, and so on. This approach seems to be universal and, probably, more transparent than the recursive computation demonstrated in section 4.

$$c^T = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & \frac{1}{8} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} & \frac{1}{16} & \frac{1}{8} & \frac{1}{16} & \frac{1}{8} & \frac{1}{16} & \frac{1}{8} & \frac{1}{16} & \frac{1}{16} & \frac{1}{32} & \frac{1}{16} & \frac{1}{32} & \frac{1}{16} & \frac{1}{32} & \frac{1}{16} & \frac{1}{32} & \frac{1}{16} & \frac{1}{32} & \frac{1}{16} & \frac{1}{32} & \frac{1}{16} & \frac{1}{32} & \frac{1}{16} & \frac{1}{32} \end{bmatrix}$$

$$A = \begin{bmatrix}
1 & . & 1 & . & . & . & 1 & . & . & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
1 & . & 1 & . & . & . & 1 & . & . & . & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
1 & . & 1 & . & . & . & . & 1 & . & . & . & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . & . & . & . \\
1 & . & 1 & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . & . & . \\
1 & . & . & 1 & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . & . \\
1 & . & . & 1 & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . \\
1 & . & . & 1 & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . \\
1 & . & . & 1 & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . \\
. & 1 & . & . & 1 & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . & . & 1 & . & . & . & . & . & . & . \\
. & 1 & . & . & 1 & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . & . & . & 1 & . & . & . & . & . & . \\
. & 1 & . & . & 1 & . & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . & . & . & 1 & . & . & . & . & . \\
. & 1 & . & . & 1 & . & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . & . & . & . & 1 & . & . & . & . \\
. & 1 & . & . & . & 1 & . & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . & . & . & . & 1 & . & . & . \\
. & 1 & . & . & . & 1 & . & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & 1 & . & . \\
. & 1 & . & . & . & 1 & . & . & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & 1 & . \\
. & 1 & . & . & . & 1 & . & . & . & . & . & . & . & 1 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & 1
\end{bmatrix} \qquad b_1 = \begin{bmatrix} s \\ s \\ s \\ s \\ s \\ s \\ s \\ s \\ s \\ s \\ s \\ s \\ s \\ s \\ s \\ s \end{bmatrix}$$

$$b_2 = \begin{bmatrix} 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 \end{bmatrix}$$

Figure 19: Components of linear program (43) for $n = 4,\ \xi_i(\omega) \in \{1, 2\},\ p_1 = p_2 = 0.5$

But, as one may guess, there are also the following disadvantages in this approach:

- since the problem size grows exponentially with the number of stages, the straightforward solution of a problem with an LP solver becomes even with a moderate number of stages unrealistic. For example, having 10 stages in the sequence and the demand state space $\{1, 2, 3\}$, we are faced with the linear program (44) of dimension ca. $90,000 \times 60,000$.
- the above approach does the job only for one fixed value of s at a time, while it were much better for us to have the expected overall performance expressed as a function of s — as we had it resolving $E|n|(s)$ recursively.

The first of these two issues can still be resolved efficiently. The special structure of our large-scale linear program makes it tractable by stochastic programming *decomposition methods* like the most frequently used *L-shaped method* [8, p. 155]. See more on this matter in the textbooks [40, 24, 8].

But still, the ability to solve our multistage problem efficiently for any given s doesn't resolve the second issue. How can we work it through?

As we could already notice in section 4, the expected recourse functions $Q(x_1)$ as well as the expected overall performance functions $E|n|(s)$ were all piecewise linear and concave. In fact, this property holds for any finite demand distributions and any number of stages [41, pp. 76, 96], [8, p. 129]. Hence, in order to express a function $E|n|(s)$ we would have to compute its values at the breakpoints by solving the corresponding stochastic program (or its deterministic equivalent) several times, and then drawing a linear spline through these points. The question is: how can we know in advance the breakpoints of function $E|n|(s)$? If we take a look on the functions $E\mathcal{M}|1|1|(s)$ and $E\mathcal{M}|1|^3(s)$ (formulas (34) and (38) as well as Figures 9 and 11 on pages 35 and 38, respectively) we see that they have their breakpoints at $s = 0, 1, 2, \ldots$. So we could guess that in case of Bernoulli distribution the breakpoints are represented by the integer values of $s = 0, \ldots, n$. But — what are they in the general case of any finite distribution?

We can refer to Powell and Cheung [34] for another approach. They consider a special class of multistage stochastic programs with the so-called *tree recourse* and develop a method that allows to find the expected recourse function exactly using a backwards recursion. This method proves to be very efficient, being able to find the expected recourse function explicitly in just seconds, even for large problems. Although our stochastic multistage program can be represented as a multistage tree recourse problem, still such special class of problems seems to be too general for us, either. Nevertheless, the approach of Powell and Cheung suggests an idea for us to construct a backwards recursion method for finding the expected recourse functions and, thus, expected overall performance functions explicitly. As we will see, this method generalizes the approach presented in the section 4 for Bernoulli demand distributions.

5.2 Backwards recursion method

As always, we assume the customer demands to be independent and identically distributed, and assume (40) to be their probability mass function. Let us consider here once again the first-stage decision making in a sequence with n stages, presented by dynamic formulation (21)–(23) in section 3.4 on page 26:

$$\begin{aligned} Z_n(\xi_1, s) &= \max \frac{x_1}{\xi_1} + Q_{n-1}(x_1) \\ x_1 &\leq s \\ x_1 &\leq \xi_1 \\ x_1 &\geq 0 \end{aligned} \tag{45}$$

where the expected recourse function $Q_{n-1}(x_1)$ absorbs all consequent stages $2, \ldots, n$:

$$\begin{aligned} Z_{n-k+1}(\xi_k, \tilde{s}) &= \max \frac{x_k}{\xi_k} + Q_{n-k}(x_k) \\ x_k &\leq \tilde{s} \\ x_k &\leq \xi_k \\ x_k &\geq 0 \\ \text{where} & \\ Q_{n-k}(x_k) &= E\left[Q_{n-k}(x_k, \xi_{k+1})\right] \\ Q_{n-k}(x_k, \xi_{k+1}) &= Z_{n-k}(\xi_{k+1}, \tilde{s} - x_k) \\ \text{with} & \\ Q_0 &\equiv 0 \end{aligned} \tag{46}$$

As we already know,[e] the expectation of the optimal objective value in (45) delivers the expected overall performance in the whole sequence provided the initial stock equals s:

$$E|n|(s) = E\left[Z_n(\xi_1, s)\right]$$

The same holds for all consequent stages: if the available stock at the beginning of a k-th stage amounts to $\tilde{s}$, then

$$E|n-k+1|(\tilde{s}) = E\left[Z_{n-k+1}(\xi_k, \tilde{s})\right]$$

[e]See Proposition 1 on page 27.

Hence, we can rewrite (46) in the following recursive form:

$$
\begin{aligned}
Z_n(\xi_1, s) \;&=\; \max \frac{x_1}{\xi_1} \;+\; Q_{n-1}(x_1) \\
x_1 \;&\leq\; s \\
x_1 \;&\leq\; \xi_1 \\
x_1 \;&\geq\; 0 \\
\text{where} \quad & \\
Q_{n-1}(x_1) \;&=\; E|n-1|(s - x_1) \\
\text{and} \quad & \\
Q_0 &\equiv 0
\end{aligned}
\qquad (\forall\, n > 0\,) \qquad (47)
$$

Thus, a recursive computation of $E|n|(s)$ can be organized as follows:

1. start with $n = 1$:
   - for each realization of random demand ξ_1 express the optimal decision x_1^* as well as the optimal objective value $Z_1(\xi_1, s)$;
   - express the expectation $E|1|(s)$;
2. proceed to $n = 2$:
   - for each realization of random demand ξ_1 express the optimal decision x_1^* as well as the optimal objective value $Z_2(\xi_1, s)$;
   - express its expectation $E|2|(s)$;

. . . .

in general, having obtained $E|n-1|(s)$ in the previous iteration,

- express the optimal decision x_1^* as well as the optimal objective value $Z_n(\xi_1, s)$;
- express the expectation $E|n|(s)$;

Let us implement this computation. We assume at first for the sake of simplicity that all demand realizations are non-zero:

$$0 < d_1 < \ldots < d_m.$$

We start with

<u>**Iteration 1**</u> **: n = 1**

Obviously, for each demand realization ξ_1 we make a decision x_1 by solving the following problem:

$$
\begin{aligned}
Z_1(\xi_1, s) \;&=\; \max \frac{x_1}{\xi_1} \\
x_1 \;&\leq\; s \\
x_1 \;&\leq\; \xi_1 \\
x_1 \;&\geq\; 0
\end{aligned}
$$

Obviously, the greatest possible value of x_1 maximizes the objective function:

$$x_1^* \;=\; \min\{\xi_1, s\}.$$

For each single demand realization we obtain the following optimal objective values depending on interrelationship between ξ_1 and s:

$$\xi_1 = d_1 \quad \Longrightarrow \quad Z_1(d_1, s) = \begin{cases} 1, & s > d_1 \\ \dfrac{s}{d_1}, & s \le d_1 \end{cases}$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$\xi_1 = d_m \quad \Longrightarrow \quad Z_1(d_m, s) = \begin{cases} 1, & s > d_m \\ \dfrac{s}{d_m}, & s \le d_m \end{cases}$$

Then the expectation of the optimal objective value with respect to the distribution of ξ_1 expresses as:

$$E|1|(s) = E_{\xi_1}\left[Z_1(\xi_1, s)\right] \;=\; p_1 \cdot Z_1(d_1, s) + \ldots + p_m \cdot Z_1(d_m, s) \;=$$

$$= \begin{cases} \displaystyle\sum_{i=1}^{m} \frac{s}{d_i} \cdot p_i, & 0 \le s \le d_1 \\ \displaystyle p_1 \;+\; \sum_{i=2}^{m} \frac{s}{d_i} \cdot p_i, & d_1 < s \le d_2 \\ \displaystyle p_1 + p_2 \;+\; \sum_{i=3}^{m} \frac{s}{d_i} \cdot p_i, & d_2 < s \le d_3 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \dots\dots\dots\dots\dots\dots \\ \displaystyle\sum_{i=1}^{m-k} p_i \;+\; \sum_{i=m-k+1}^{m} \frac{s}{d_i} \cdot p_i, & d_{m-k} < s \le d_{m-k+1} \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \dots\dots\dots\dots\dots\dots \\ \displaystyle\sum_{i=1}^{m-2} p_i \;+\; \frac{s}{d_{m-1}} \cdot p_{m-1} \;+\; \frac{s}{d_m} \cdot p_m, & d_{m-2} < s \le d_{m-1} \\ \displaystyle\sum_{i=1}^{m-1} p_i \;+\; \frac{s}{d_m} \cdot p_m, & d_{m-1} < s \le d_m \\ \displaystyle\sum_{i=1}^{m} p_i = 1, & d_m < s \end{cases} \tag{48}$$

Figure 20 depicts an example of the function $E|1|(s)$ for the uniform discrete distribution on $\{1, \ldots, 5\}$.

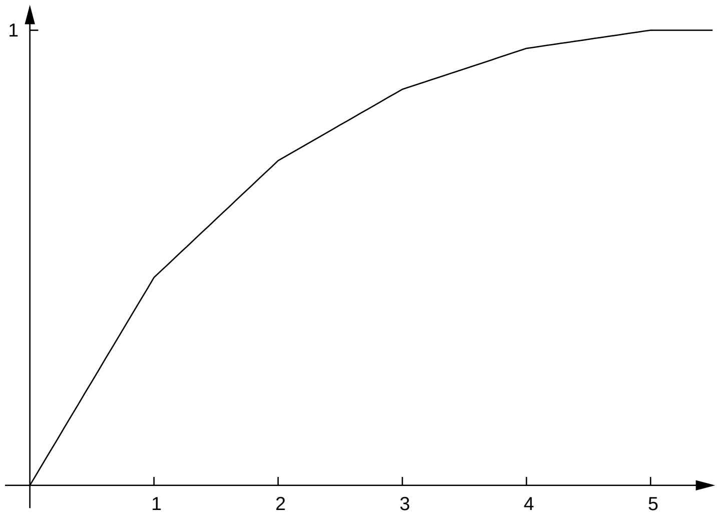


Figure 20: Function $E|1|(s)$ for ξ_1 distributed uniformly on $\{1, \ldots, 5\}$

Remark: We have to stress here that $E|1|(s)$ exemplifies the following properties:

(a) it is piecewise linear;

(b) concave;

(c) nondecreasing;

(d) its rate of increase is by itself nonincreasing: the linear pieces of the function are sloping less and less as s grows;

(e) it attains its maximal value of $n = 1$ at some point and stays constant after this rightmost breakpoint.

These properties of the function (48) can be verified explicitly. Also, piecewise linearity and concavity can be derived from the general properties of expected recourse functions mentioned on page 47.

This completes the 1st iteration. Let us now describe the n-th iteration, for any $n > 1$.

Iteration n : $\mathbf{n > 1}$

We now consider a sequence of n customers. Decision making in its 1st stage is represented by the problem (47). It is assumed that we have obtained function $E|n-1|(s)$ on the previous iteration. Let us use induction in this iteration: we assume $E|n-1|(s)$ to exemplify the properties **(a)**–**(e)** and write it down as:

$$E|n-1|(s) \;=\; \begin{cases} a_1 s + b_1, & 0 = c_0 \le s \le c_1 \\ a_2 s + b_2, & c_1 < s \le c_2 \\ \dots\dots & \dots\dots\dots\dots \\ a_r s + b_r, & c_{r-1} < s \le c_r \\ n-1, & c_r < s \end{cases} \tag{49}$$

with $0 < a_r < \ldots < a_1$.

Property **(e)** should be reformulated in the following way:

(e) it attains its maximal value of $n-1$ at some point and stays constant after this rightmost breakpoint.

We will show that the same properties also hold for $E|n|(s)$, which we compute in the current iteration. The 1st iteration proved already $E|\mathbf{1}|(s)$ to have these properties.

In order to express $E|n|(s)$, let us determine the optimal decision x_1^* for each possible realization of demand ξ_1. Obviously, we can rewrite (47) as a problem of maximization of the function of a single variable on a closed interval:

$$Z_n(\xi_1, s) \;=\; \max_{[0,\min\{\xi_1,s\}]} \frac{x_1}{\xi_1} \;+\; Q_{n-1}(x_1) \tag{50}$$

where

$$Q_{n-1}(x_1) \;=\; E|n-1|(s-x_1) \;=\; \begin{cases} a_1(s-x_1)+b_1, & c_0 \le s-x_1 \le c_1 \\ a_2(s-x_1)+b_2, & c_1 < s-x_1 \le c_2 \\ \dots\dots\dots\dots & \dots\dots\dots\dots\dots \\ a_r(s-x_1)+b_r, & c_{r-1} < s-x_1 \le c_r \\ n-1, & c_r < s-x_1 \end{cases} \;=$$

$$= \begin{cases} -a_1x_1 + a_1s + b_1, & s-c_1 \le x_1 \le s-c_0 \\ -a_2x_1 + a_2s + b_2, & s-c_2 \le x_1 < s-c_1 \\ \dots\dots\dots\dots & \dots\dots\dots\dots\dots\dots \\ -a_rx_1 + a_rs + b_r, & s-c_r \le x_1 < s-c_{r-1} \\ n-1, & x_1 < s-c_r \end{cases} = \begin{cases} n-1, & x_1 < s-c_r \\ -a_rx_1 + a_rs + b_r, & s-c_r \le x_1 < s-c_{r-1} \\ \dots\dots\dots\dots & \dots\dots\dots\dots\dots\dots \\ -a_2x_1 + a_2s + b_2, & s-c_2 \le x_1 < s-c_1 \\ -a_1x_1 + a_1s + b_1, & s-c_1 \le x_1 \le s-c_0 \end{cases} \tag{51}$$

Note that properties **(a)**–**(b)** hold for $Q_{n-1}(x_1)$ too, while **(c)**–**(e)** change to:

(c') it is nonincreasing;

(d') its rate of decrease is by itself nondecreasing: the linear pieces of the function become steeper as x_1 grows;

(e') its maximal value $n-1$ stays constant up until reaching the first breakpoint from the left.

Figure 21 shows an example of $Q_{n-1}(x_1)$ for $n = 2$ and some s, based on the example given in Figure 20.

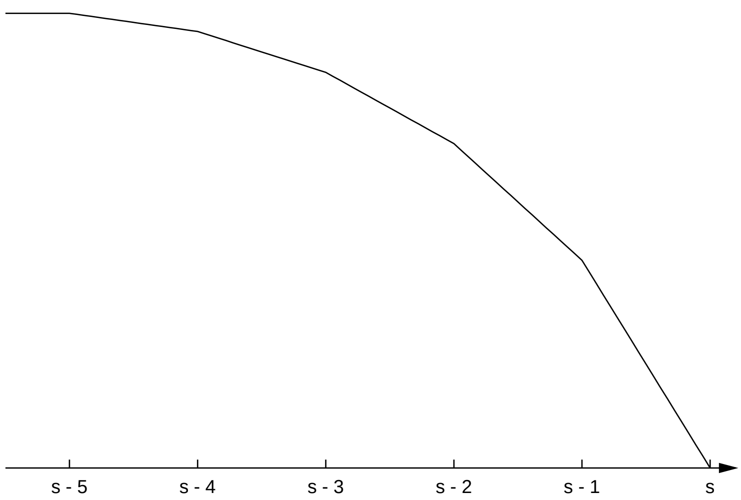


Figure 21: Function $Q_{n-1}(x_1)$ for $n = 2$ and some s ; see also Figure 20

Note that the shape of the graph of $Q_{n-1}(x_1)$ doesn't depend on s. The varying value of s makes the graph only drift along the horizontal axis. Let us enumerate the intervals of linearity of Q_{n-1} from the right to the left, so that the 1st interval corresponds to the slope $-a_1$, the 2nd one — to $-a_2$, and so on. Note that all intervals comprise together the domain $(-\infty, s]$ of the function Q_{n-1}.

Let us now turn to the objective function of (50) and denote it by $f(x_1)$:

$$f(x_1) \;=\; \frac{x_1}{\xi_1} \;+\; Q_{n-1}(x_1).$$

Obviously, its first term is a linear function, while the second one is a piecewise linear concave function. Hence, $f(x_1)$ is piecewise linear and concave, too. Depending on the value of ξ_1 and on the slopes of $Q_{n-1}(x_1)$, f may occur to be either nondecreasing, or increasing up to some point and then decreasing. The latter case is shown in Figure 22.

It is not hard to determine the value $\hat{x}_1$ that delivers a global maximum to the function f: starting at its leftmost linear piece, we have to move to the right as long as the slopes of successive pieces stay positive. Let us keep for f the enumeration of intervals of linearity introduced above for Q. Then, the slope of the i-th linear piece of f equals

$$\frac{1}{\xi_1} - a_i.$$

We know that $a_r < \ldots < a_1$; hence, the right-hand end of the rightmost interval for which holds

$$\frac{1}{\xi_1} - a_i > 0,$$

maximizes the objective function on its whole domain $(-\infty, s]$. If we denote by i^* the index of the corresponding interval then we can write down the following:

$$i^* = \min\left\{ i = 1, \ldots, r \;\middle|\; \frac{1}{\xi_1} - a_i > 0 \right\} \qquad \Longrightarrow \qquad \hat{x}_1 = s - c_{i^*-1} \tag{52}$$

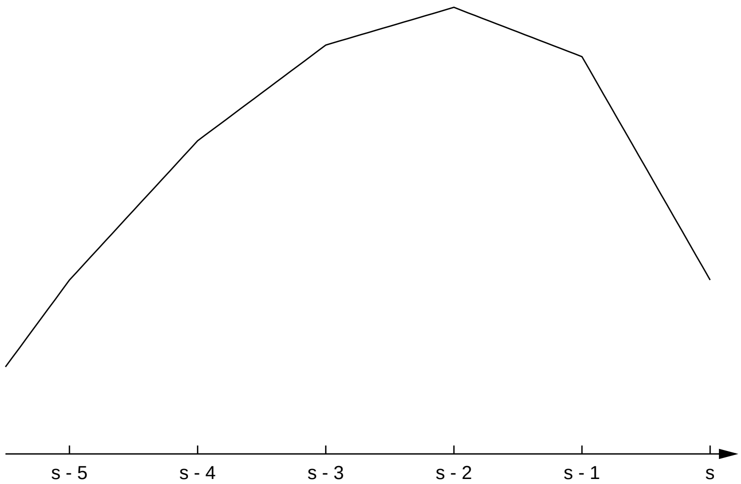


Figure 22: An example of objective function $f(x_1)$; see also Figure 21

Note that there always exists such $\hat{x}_1$ since the starting (leftmost) slope of the objective function is always positive.

For the sake of brevity we would like to simplify slightly the notation just introduced:

$$\text{we let} \qquad t = i^* - 1, \qquad \text{then:} \qquad \hat{x}_1 = s - c_t. \tag{53}$$

(index t stands for "top").

It is important to notice here that the above search for $\hat{x}_1$ is invariant of s, since the slopes of $f(x_1)$ are invariant of s. Hence, for each s it will always be the same linear piece of a polygon line with its rightmost point at the top of the graph. Figure 23 demonstrates this by drawing function $f(x_1)$ for several values of parameter s.

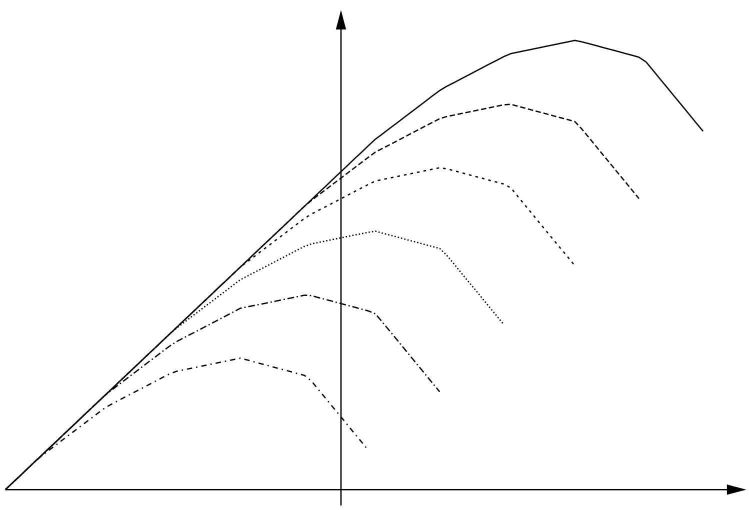


Figure 23: Instances of function $f(x_1)$ for several values of parameter s

But, since we are maximizing $f(x_1)$ on the interval $[0, \min\{\xi_1, s\}]$, the global maximizer $\hat{x}_1$ may occur to be infeasible. We need to be able to determine a feasible maximizer x_1^* as well as to express the corresponding optimal objective value. Doing this, it is necessary to distinguish between the following three cases, which we sketch in Figure 24.

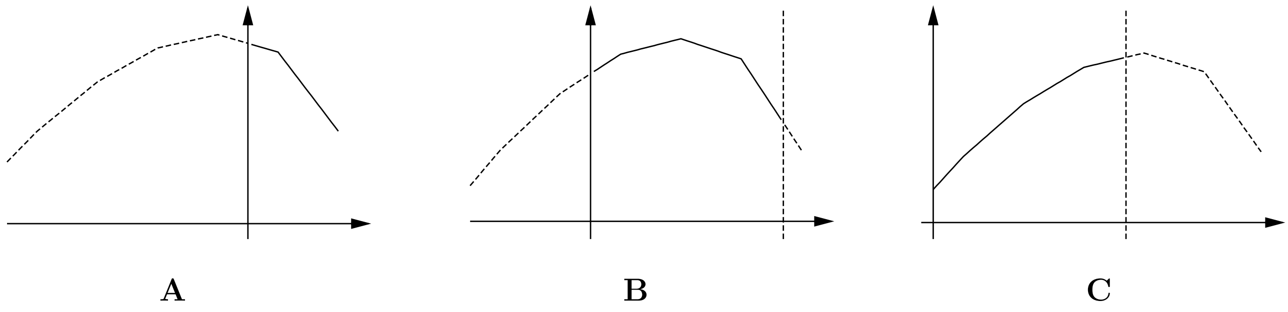


Figure 24: Determining a feasible maximizer

Case A: This is the case when the topmost point of the polygon is located behind the vertical axis, and so the function doesn't have any positive slope while getting into domain $x_1 > 0$. This situation arises by sufficiently low values of s and is characterized by the inequality

$$\hat{x}_1 = s - c_t \leq 0.$$

This forces the optimal solution to take the lowest possible value:

$$x_1^* = 0.$$

Let us express the corresponding optimal objective value in this case:

$$\begin{aligned}
Z_n(\xi_1, s) \;=\; f(x_1^*) \;=\; \frac{0}{\xi_1} + Q_{n-1}(0) \;= \\
= \begin{cases} a_t s + b_t, & s - c_t \leq 0 < s - c_{t-1} \\ \cdots\cdots & \cdots\cdots\cdots\cdots \\ a_2 s + b_2, & s - c_2 \leq 0 < s - c_1 \\ a_1 s + b_1, & s - c_1 \leq 0 \leq s - c_0 \end{cases} = \begin{cases} a_1 s + b_1, & c_0 \leq s \leq c_1 \\ a_2 s + b_2, & c_1 < s \leq c_2 \\ \cdots\cdots & \cdots\cdots \\ a_t s + b_t, & c_{t-1} < s \leq c_t \end{cases}
\end{aligned} \tag{54}$$

Obviously, the number of cases in (54) is determined by the number of edges in the polygon *after* its topmost point. The two lowest polygons in Figure 23 clarify this: depending on how far $\hat{x}_1$ lies behind the vertical axis, $x_1^* = 0$ can be found belonging to any of the declining edges, i.e., to any of the intervals after $\hat{x}_1 = s - c_t$. This requires from us to use each time an appropriate linear piece for evaluation of the objective function at 0.

Case B: In this case, there is a feasible topmost vertex between 0 and $\min\{\xi_1, s\}$, i.e., this case is characterized by the inequality

$$0 < \hat{x}_1 \leq \xi_1 \qquad \text{or} \qquad 0 < s - c_t \leq \xi_1$$

Hence,

$$x_1^* = \hat{x}_1 = s - c_t$$

Also, we know definitely which linear piece to use in order to evaluate f at this optimal point. Thus we don't need to branch the optimal objective value, as we did in case A. So we obtain:

$$\begin{aligned} Z_n(\xi_1, s) &= f(x_1^*) = \frac{s - c_t}{\xi_1} + Q_{n-1}(s - c_t) = \\ &= \frac{s - c_t}{\xi_1} + a_t(s - (s - c_t)) + b_t = \frac{s}{\xi_1} - \frac{c_t}{\xi_1} + a_t c_t + b_t. \end{aligned} \tag{55}$$

Case C: This is the case when a sufficiently low demand realization ξ_1 makes $\hat{x}_1$ infeasible:

$$\xi_1 < \hat{x}_1 = s - c_t.$$

This forces the optimal solution to take the greatest possible value:

$$x_1^* = \xi_1,$$

since the function is monotone increasing from the left of $\hat{x}_1$.

Let us express the corresponding optimal objective value. Again, as in case A, we have to branch the expression as many times, as many linear pieces does the polygon contain *before* its topmost point, since the cut ξ_1 can become realized at each of these edges — depending on the value of s. This requires from us to use each time an appropriate linear piece for evaluation of the objective function at $x_1^* = \xi_1$.

$$\begin{aligned} Z_n(\xi_1, s) &= f(x_1^*) = \frac{\xi_1}{\xi_1} + Q_{n-1}(\xi_1) = \\ &= 1 + \begin{cases} n - 1, & \xi_1 < s - c_r \\ -a_r\xi_1 + a_r s + b_r, & s - c_r \le \xi_1 < s - c_{r-1} \\ \dots\dots\dots\dots\dots & \dots\dots\dots\dots\dots \\ -a_{t+1}\xi_1 + a_{t+1}s + b_{t+1}, & s - c_{t+1} \le \xi_1 < s - c_t \end{cases} = \\ &= \begin{cases} a_{t+1}(s - \xi_1) + b_{t+1} + 1, & c_t + \xi_1 < s \le c_{t+1} + \xi_1 \\ \dots\dots\dots\dots\dots & \dots\dots\dots\dots\dots \\ a_r(s - \xi_1) + b_r + 1, & c_{r-1} + \xi_1 < s \le c_r + \xi_1 \\ n, & c_r + \xi_1 < s \end{cases} \end{aligned} \tag{56}$$

Let us now unify the above three cases. As we have seen, they are characterized by the 3 inequalities, which we have rewritten here to the right:

$$\begin{cases} s - c_t \le 0 \\ 0 < s - c_t \le \xi_1 \\ \xi_1 < s - c_t \end{cases} \implies \begin{cases} s \le c_t \\ c_t < s \le c_t + \xi_1 \\ c_t + \xi_1 < s \end{cases}$$

Hence, with the above three cases we compute an optimal decision and the optimal objective value $Z_n(\xi_1, s)$ for each possible value of s, under given demand realization ξ_1. Moreover, the expressions (54), (55), (56) of $Z_n(\xi_1, s)$ can be unified into the single one:

$$Z_n(\xi_1, s) \;=\; \begin{cases} a_1 s + b_1, & 0 = c_0 \le s \le c_1 \\ a_2 s + b_2, & c_1 < s \le c_2 \\ \dots\dots\dots\dots\dots\dots & \dots\dots\dots\dots\dots\dots \\ a_t s + b_t, & c_{t-1} < s \le c_t \\ \dfrac{s}{\xi_1} - \dfrac{c_t}{\xi_1} + a_t c_t + b_t, & c_t < s \le c_t + \xi_1 \\ a_{t+1}(s - \xi_1) + b_{t+1} + 1, & c_t + \xi_1 < s \le c_{t+1} + \xi_1 \\ \dots\dots\dots\dots\dots\dots & \dots\dots\dots\dots\dots\dots \\ a_r(s - \xi_1) + b_r + 1, & c_{r-1} + \xi_1 < s \le c_r + \xi_1 \\ n, & c_r + \xi_1 < s \end{cases} \tag{57}$$

Let us show that $Z_n(\xi_1, s)$ is a function in s that satisfies the properties **(a)–(e)**, which we assumed to hold for $E|n-1|(s)$ at the beginning of the iteration:

(a) Piecewise linearity: obvious.

(b) Concavity. First of all, we have to check the continuity. Continuity of $Z_n(\xi_1, s)$ in three separate parts (54), (55), (56) follows from continuity of $E|n-1|(s)$ given by (49); an explicit verification shows that the function is also continuous at the points c_t and $c_t + \xi_1$. Then, its concavity becomes delivered by the properties **(c)–(d)**:

(c) Nondecreasing: the slopes of all linear pieces are nonnegative, since $0 < a_r < \ldots < a_1$ and $\xi_1 > 0$.

(d) Its rate of increase is by itself nonincreasing: the linear pieces of the function are sloping less and less as s grows. Indeed, the same property of $E|n-1|(s)$ provides us with $a_r < \ldots < a_1$, and hence the property holds for $s < c_t$ and $s > c_t + \xi_1$. Let us show that it holds also at c_t and $c_t + \xi_1$, i.e., that the function doesn't become steeper at any of these two points. According to (52)–(53), index t is the lowest one that fulfills

$$\frac{1}{\xi_1} - a_{t+1} > 0,$$

what means that for the lower index $t-1$ can only hold

$$\frac{1}{\xi_1} - a_{(t-1)+1} \le 0.$$

Rewriting the above inequalities as

$$a_{t+1} < \frac{1}{\xi_1} \qquad \text{and} \qquad \frac{1}{\xi_1} \le a_t,$$

we conclude that the slopes at the above two breakpoints do not increase, indeed.

(e) it attains its maximal value of $n = 1$ at some point and stays constant after this rightmost breakpoint: obvious.

Expressing the recourse functions $Z_n(\xi_1, s)$ for each realization of $\xi_1 = d_1, \ldots, d_m$, we finally come to the expected recourse function, which represents the expected overall performance:

$$E|n|(s) \;=\; E\left[Z_n(\xi_1, s)\right] \;=\; p_1 \cdot Z_n(d_1, s) + \ldots + p_m \cdot Z_n(d_m, s)$$

Since we produce a convex combination of the functions satisfying properties **(a)**–**(e)**, the latter are kept by $E|n|(s)$, too. This proves the induction and completes the iteration. □

The above recursion method allows us to compute the expected recursion function for any n explicitly. Comparing the non-mobile configuration $\mathrm{C}|n_1|\ldots|n_K|$ of a distribution system with its mobile counterpart $\mathrm{C}\mathcal{M}|n_1|\ldots|n_K|$ in terms of their expected overall performances

$$E|n_1|\ldots|n_K|(s) = E|n_1|(sn_1/N) + \ldots + E|n_K|(sn_K/N)$$

and

$$E\mathcal{M}|n_1|\ldots|n_K|(s) = E|N|(s)$$

we have to compute the functions $E|n_1|, \ldots, E|n_K|, E|N|$. Running the recursive method through $1, \ldots, N$, we obtain these functions, and hence, can compare the performance in mobile and non-mobile environments. The next sections presents some computational experience we have got for several different configurations of a distribution system.

Remark 1: We presented the method keeping for the sake of simplicity the assumption of non-zero demands: $0 < d_1 < \ldots < d_m$. It is not hard to adjust the method also for the case of having a zero demand in the probability distribution.

Remark 2: As we can also see, the method doesn't restrict us to one and the same demand distribution for all customers, but allows each customer demand to follow its own probability distribution, which only has to be a finite one. Nevertheless, we restrict us in this paper to the case of identically distributed demands.

Remark 3: Approximate solutions for the case of continuous demand distributions can also be obtained with the presented method, since we can approximate continuous distributions by discrete ones through sampling of the former.

5.3 Computational results

Example 1

Let us consider the non-mobile configuration $\mathrm{C}|5|5|5|5|5| \equiv \mathrm{C}|5|^5$ of a distribution system with 25 customers, and compare its expected overall performance with that of the mobile configuration $\mathrm{C}\mathcal{M}|5|^5$. We assume the demand distribution to be dicrete uniform with the state space $\{1, 2, \ldots, 10\}$.

Figure 25 presents the expected performance functions and the corresponding advantage of the mobile solution in per cent. It took the computer with an AMD Sempron™ 1.83 GHz processor less than a second to obtain the $E|25|(s)$ function. The latter turns out to be piecewise linear consisting of 251 pieces.

As we can see on the advantage graph (to the right), the relative advantage increases for very scarce resources: we had to truncate the graph for values of $s < 20$ with performance advantage over

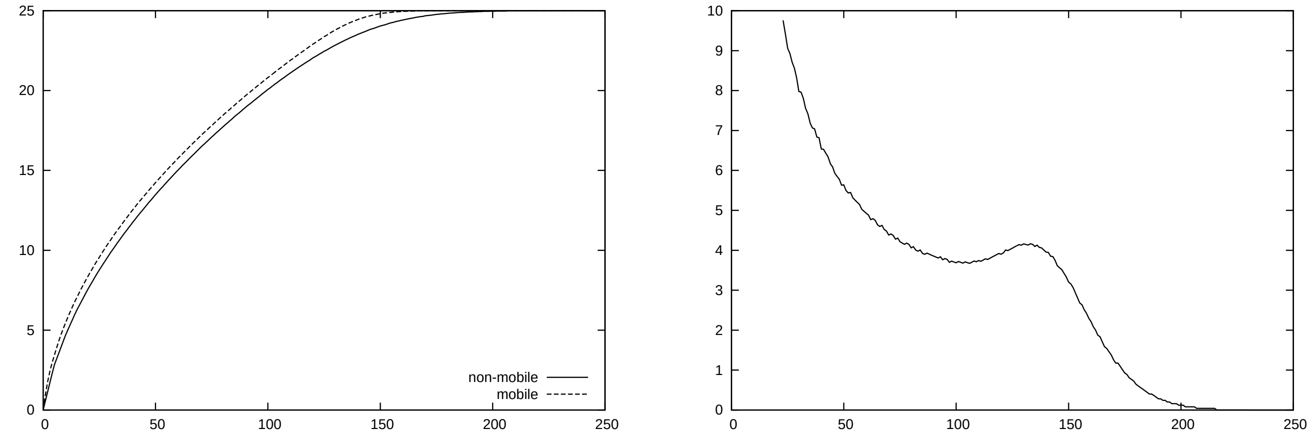


Figure 25: Expected performances $E|\mathsf{5}|^{\mathsf{5}}(s)$ and $E\mathcal{M}|\mathsf{5}|^{\mathsf{5}}(s)$, and the relative advantage of mobile performance (in %)

10%, either, in order to keep an appropriate scaling for the rest of the graph. A peak performance improvement around average values of s is 4.16% attained at $s = 132$.

Example 2

The above peak performance improves slightly if we consider a demand distribution with a greater performance. In case presented in Figure 26 we took 10 equally spaced values between 1 and 100 as equally probable demand realizations.

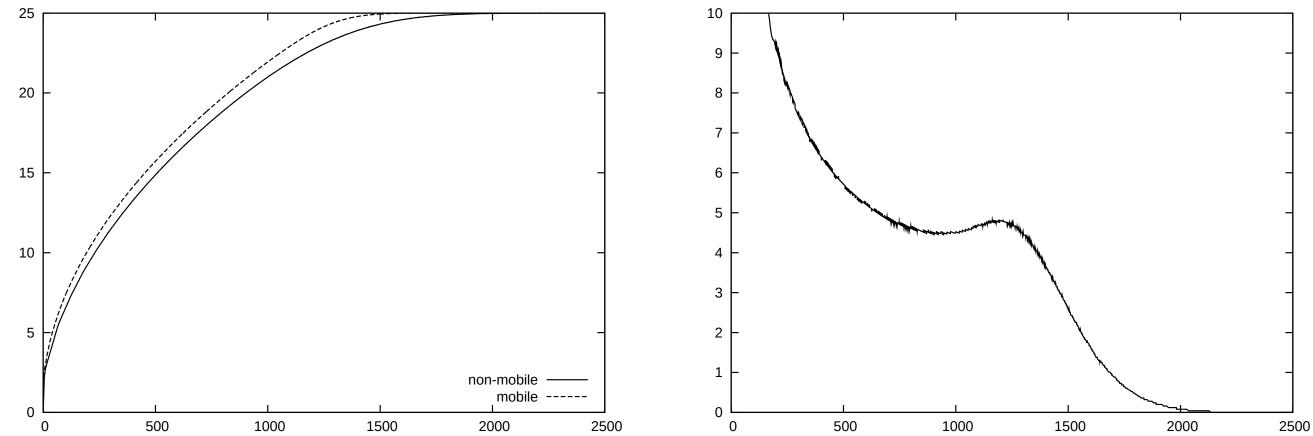


Figure 26: Expected performances $E|\mathsf{5}|^{\mathsf{5}}(s)$ and $E\mathcal{M}|\mathsf{5}|^{\mathsf{5}}(s)$, and the relative advantage of mobile performance (in %)

The similar peak performance attains in this case the value 4.79%.

Example 3

The following example shows that the improvement effect vanishes as larger customer groups become integrated in the mobile solution.

We consider here the non-mobile configuration $\mathrm{C}|\mathsf{50}|\mathsf{50}|\mathsf{50}|\mathsf{50}|\mathsf{50}| \equiv \mathrm{C}|\mathsf{50}|^{\mathsf{5}}$ of a distribution system with 250 customers, and compare its expected overall performance with that of the mobile configuration $\mathrm{C}\mathcal{M}|\mathsf{50}|^{\mathsf{5}}$. We assume the demand distribution to be the same one as in Example 1 — discrete uniform with the state space $\{1, 2, \ldots, 10\}$.

Figure 27 presents the expected performance functions and the corresponding advantage of the mobile solution in per cent. It took the same computer 30 seconds to obtain $E|\mathsf{250}|(s)$. The latter turns out to be piecewise linear consisting of 2501 pieces.

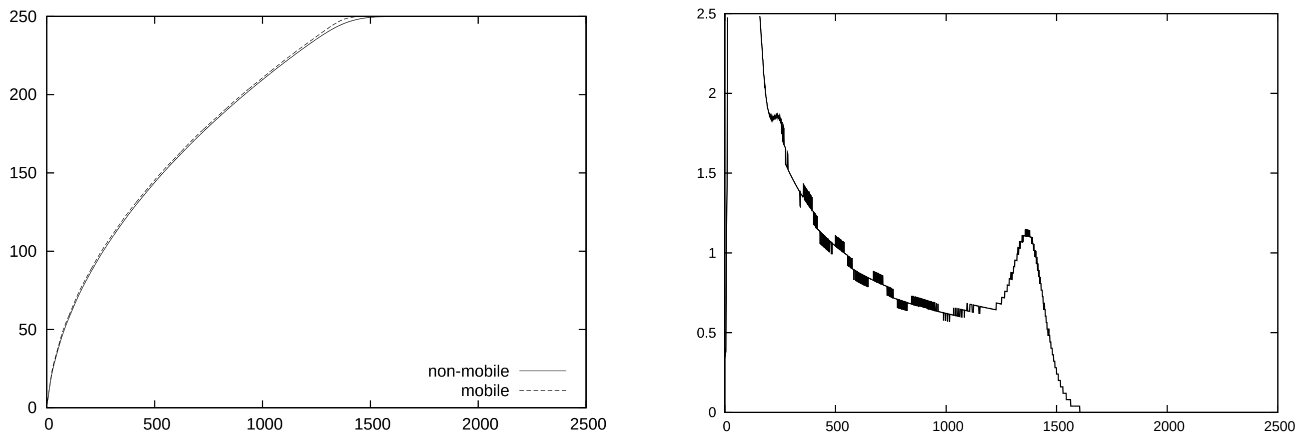


Figure 27: Expected performances $E|\mathsf{50}|^{\mathsf{5}}(s)$ and $E\mathcal{M}|\mathsf{50}|^{\mathsf{5}}(s)$, and the relative advantage of mobile performance (in %)

The similar peak performance attains in this case the value 1.13%.

A Comparing the expected performance functions

Comparing $E|1|1|(s)$ and $E\mathcal{M}|1|1|(s)$

Proposition 3 $\forall p \in (0,1) \quad \forall s \geq 0 \quad : \quad E|1|1|(s) \;\leq\; E\mathcal{M}|1|1|(s).$

Proof: We can split the proof into three parts:

(1:) $0 \leq s \leq 1$:

$$\begin{aligned} E|1|1|(s) = 2 - p(2-s) &\leq 2(1-p) + ps(2-p) = E\mathcal{M}|1|1|(s) \\ 2 - 2p + ps &\leq 2 - 2p + 2ps - p^2 s \\ 0 &\leq ps - p^2 s \\ 0 &\leq ps(1-p) \end{aligned}$$

(2:) $1 < s \leq 2$:

$$\begin{aligned} E|1|1|(s) = 2 - p(2-s) &\leq 2 - p^2(2-s) = E\mathcal{M}|1|1|(s) \\ -p(2-s) &\leq -p^2(2-s) \\ p(2-s) &\geq p^2(2-s) \end{aligned}$$

(3:) $s > 2$:

$$E|1|1|(s) = 2 \;\leq\; 2 = E\mathcal{M}|1|1|(s)$$

□

Comparing $E|1|^3(s)$ and $E\mathcal{M}|1|^3(s)$

Proposition 4 $\forall p \in (0,1) \quad \forall s \geq 0 \quad : \quad E|1|^3(s) \leq E\mathcal{M}|1|^3(s).$

Proof: We can split the proof into four parts:

(1:) $0 \leq s \leq 1$:

$$\begin{aligned} E|1|^3(s) = 3 - p(3-s) &\leq 3 - 3p + ps(2-p) + ps(1-p)^2 = E\mathcal{M}|1|^3(s) \\ 3 - 3p + ps &\leq 3 - 3p + ps(2-p) + ps(1-p)^2 \\ ps &\leq ps(2-p) + ps(1-p)^2 \\ s &\leq s(2-p) + s(1-p)^2 \end{aligned}$$

The last inequality holds since $2 - p > 1$ and $(1-p)^2 > 0$.

(2:) $1 < s \leq 2$:

$$\begin{aligned} E|1|^3(s) = 3 - p(3-s) &\leq 3 - 3p^2(2-s) - p^3(2s-3) = E\mathcal{M}|1|^3(s) \\ -p(3-s) &\leq -3p^2(2-s) - p^3(2s-3) \\ 3 - s &\geq 3p(2-s) + p^2(2s-3) \end{aligned}$$

Clearly, the right-hand side is a strictly increasing function in p. Keeping in mind that $0 < p < 1$, let us assign the limiting value 1 to p. Then the right-hand side expresses as:

$$3 \cdot 1 \cdot (2 - s) + 1 \cdot (2s - 3) = 6 - 3s + 2s - 3 = 3 - s.$$

Thus, the right-hand side equals the left-hand side under $p = 1$. It follows that for $0 < p < 1$ the following inequality holds:

$$3 - s \ > \ 3p(2 - s) + p^2(2s - 3)$$

(3:) $2 < s \leq 3$:

$$\begin{array}{rcl} E|1|^3(s) = 3 - p(3 - s) & \leq & 3 - p^3(3 - s) = E\mathcal{M}|1|^3(s) \\ -p(3 - s) & \leq & -p^3(3 - s) \\ p(3 - s) & \geq & p^3(3 - s) \end{array}$$

(4:) $s > 3$:

$$E|1|^3(s) = 3 \ \leq \ 3 = E\mathcal{M}|1|^3(s)$$

$\square$

References

[1] D. Adelman. A price-directed approach to stochastic inventory/routing. *Operations Research*, 52(4):499–514, 2004.

[2] S. Anily and A. Federgruen. One warehouse multiple retailer systems with vehicle routing costs. *Management Science*, 36:92–114, 1990.

[3] S. Anily and A. Federgruen. Two-echelon distribution systems with vehicle routing costs and central inventories. *Operations Research*, 41:37–47, 1993.

[4] R. Baldacci, E. Hadjiconstantinou, and A. Mingozzi. An exact algorithm for the capacitated vehicle routing problem based on a two-commodity network flow formulation. *Operations Research*, 52(5):723–738, 2004.

[5] M. O. Ball, C.-Y. Chen, and Z.-Y. Zhao. Available to promise. *Smith Papers Online*, 2002. `http://mail3.rhsmith.umd.edu/Faculty/KM/papers.nsf`.

[6] D. J. Bertsimas. A vehicle routing problem with stochastic demand. *Operations Research*, 40:574–585, 1992.

[7] D. J. Bertsimas and D. Simchi-Levi. A new generation of vehicle routing research: robust algorithms, addressing uncertainty. *Operations Research*, 44:286–304, 1996.

[8] J. R. Birge and F. Louveaux. *Introduction to Stochastic Programming*. Springer-Verlag, New York, 1997.

[9] J. Bramel and D. Simchi-Levi. A location based heuristic for general routing problems. *Operations Research*, 43:649–660, 1995.

[10] T. Carvalho and W. Powell. A multiplier adjustment method for dynamic resource allocation problems. *Transportation Science*, 34(2):150–164, 2000.

[11] C.-Y. Chen, Z.-Y. Zhao, and M. O. Ball. Quantity and due date quoting available to promise. *Information Systems Frontiers*, 3(4):477–488, 2001.

[12] R. K.-M. Cheung and W. B. Powell. Models and algorithms for distribution problems with uncertain demands. *Transportation Science*, 30(1):43–59, 1996.

[13] R.-M. Cheung and W. Powell. An algorithm for multistage dynamic networks with random arc capacities, with an application to dynamic fleet management. *Operations Research*, 44(6):951–963, 1996.

[14] M. Dror, G. Laporte, and P. Trudeau. Vehicle routing with stochastic demands: properties and solution frameworks. *Transportation Science*, 23:166–176, 1989.

[15] Employment Development Department, State of California, USA. *California Occupational Guide*. `http://www.calmis.cahwnet.gov/file/occguide`.

[16] A. Federgruen, G. Prastacos, and P. Zipkin. An allocation and distribution model for perishable products. *Operations Research*, 34(1):75–82, 1986.

[17] A. Federgruen and P. Zipkin. A combined vehicle routing and inventory allocation problem. *Operations Research*, 32(5):1019–1037, 1984.

[18] L. Frantzeskakis and W. B. Powell. A successive linear approximation procedure for stochastic, dynamic vehicle allocation problems. *Transportation Science*, 24(1):40–57, 1990.

[19] M. Gendreau, G. Laporte, and R. Segun. A tabu search heuristic for the vehicle routing problem with stochastic demands and customers. *Operations Research*, 44:469–477, 1996.

[20] G. Godfrey and W. B. Powell. An adaptive, distribution-free algorithm for the newsvendor problem with censored demands, with application to inventory and distribution problems. *Management Science*, 47(8):1101–1112, 2001.

[21] E. Goldberg. Drilling down and bubbling up. *BeverageWorld Supplement*, 44:22–26, May 1993.

[22] B. L. Golden and A. A. Assad, editors. *Vehicle Routing: Methods and Studies*. North-Holland, Amsterdam, The Netherlands, 1988.

[23] P. Kall. *Stochastic linear programming*. Springer-Verlag, Berlin Heidelberg New York, 1976.

[24] P. Kall and S. W. Wallace. *Stochastic Programming*. John Wiley & Sons, Chichester, 1994.

[25] A. J. Kleywegt, V. S. Nori, and M. W. P. Savelsbergh. The stochastic inventory routing problem with direct deliveries. *Operations Research*, 52(4):499–514, 2004.

[26] G. Laporte. The vehicle routing problem: An overview of exact and approximate algorithms. *European Journal of Operational Research*, 59:345–358, 1992.

[27] G. Laporte, F. V. Louveaux, and L. van Hamme. An integer L-shaped algorithm for the capacitated vehicle routing problem with stochastic demands. *Operations Research*, 50(3):415–423, 2002.

[28] M. W. Lund. Real options in offshore oil field development projects. In *3rd Annual Real Options Conference*, Leiden, 1999. `http://www.realoptions.org/papers1999/LUND.PDF`, Download 2004-12-07.

[29] A. Maatman, C. Schweigman, A. Ruijs, and M. H. van der Vlerk. Modeling farmer's response to uncertain rainfall in Burkina Faso: a stochastic programming approach. *Operations Research*, 50(3):399–414, 2002.

[30] C. T. C. Mo, S. S. Wu, R. Chen, and M. C. K. Yang. Optimal sequential allocation with imperfect feedback information. *Journal of Applied Probability*, 38(1):248–254, 2001.

[31] Mobile Data Solutions, Richmond, BC, Canada. MDSI and MRO to offer work and asset management system to the utility industry. `http://www.mdsi.ca/about_mdsi/news/news11220401.asp`, 11 2004.

[32] W. B. Powell. *A Comparative Review of Alternative Algorithms for the Dynamic Vehicle Allocation Problem*, pages 249–291. In Golden and Assad [22], 1988.

[33] W. B. Powell. *Dynamic Asset Management*. CASTLE Laboratory, Department of Operations Research and Financial Engineering, Princeton University, USA, 2004. `http://www.castlelab.princeton.edu/`.

[34] W. B. Powell and R. K. Cheung. Stochastic programs over trees with random arc capacities. *Networks*, 24:161–175, 1994.

[35] W. B. Powell and H. Topaloglu. *Stochastic Programming in Transportation and Logistics*, chapter 9, pages 555–636. Volume 10 of Ruszczyński and Shapiro [42], 2003.

[36] Prosy GmbH, Wuppertal, Germany. Mobile logistics. `http://www.prosy.de/produkte/mobilogi/allgemein.htm` (in German).

[37] H. N. Psaraftis. Dynamic vehicle routing: Status and prospects. *Annals of Operations Research*, 61:143–164, 1995.

[38] A. V. Puelz. A stochastic convergence model for portfolio selection. *Operations Research*, 50(3):462–476, 2002.

[39] D. F. Ross. *Distribution: Planning and Control.* Chapman & Hall, New York, 1996.

[40] A. Ruszczyński. *Decomposition Methods*, chapter 3, pages 41–211. Volume 10 of Ruszczyński and Shapiro [42], 2003.

[41] A. Ruszczyński and A. Shapiro. *Optimality and Duality in Stochastic Programming*, chapter 2, pages 65–139. Volume 10 of *Handbooks in Operations Research and Management Science* [42], 2003.

[42] A. Ruszczyński and A. Shapiro, editors. *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science.* Elsevier, Amsterdam, 2003.

[43] A. Ruszczyński and A. Shapiro. *Stochastic Programming Models*, chapter 1, pages 1–64. Volume 10 of *Handbooks in Operations Research and Management Science* [42], 2003.

[44] N. Secomandi. A rollout policy for the vehicle routing problems with stochastic demands. *Operations Research*, 49:796–802, 2001.

[45] U.S. Bureau of Labor Statistics, USA. *Occupational Outlook Handbook*, 2004-05 edition. `http://www.bls.gov/oco/home.htm`.

[46] W.-H. Yang, K. Mathur, and R. H. Ballou. Stochastic vehicle routing with restocking. *Transportation Science*, 34:99–112, 2000.