Developing operational understanding of multi - channel service delivery systems through computer simulation

Isabella Porcelli (isabella.porcelli@unifi.it)
Università degli Studi di Firenze, Italy

Mario Rapaccini
Università degli Studi di Firenze, Italy

Filippo Visintin
Università degli Studi di Firenze, Italy

Abstract
The paper aims to investigate the effects of different queue management policies in service delivery through computer simulation. The effects are evaluated both on objective performance (i.e. throughput time) as well on performance perceived by customers (i.e. customers' satisfaction and perceived waiting times). The importance of establishing an integrated approach to take into account not only the management logics, but also how people interact with these systems according to their individual behaviours and attitudes, is then emphasized.

Keywords: service delivery system, perceived waiting time, computer simulation

Introduction
Several authors affirm the urgency for rigorous study to guide service managers in improving the design, efficiency, and effectiveness of Service Delivery Systems (SDSs) (Roth and Menor, 2003; Metters and Marucheck, 2007). As stated by Sampson (2000), "all services have customers as primary suppliers of inputs". The customer presence in the delivery process - i.e. *the moment of truth* (Carlzon, 1989) - implies a lot of variability to be taken into account. This can stem from individual behaviours of customers (e.g. personal traits, temporary mood, ability in using technology, etc.), and/or from specific management policies (i.e. the willingness to profile the customers that enter into the systems, in order to handle them differently). The service marketing literature, in addition, is plenty of publications that give evidences that the customer's satisfaction depends not on the actual system performance, but from those perceived by the customer with respect to its expectations (Parasuraman et al., 1994; Davis and Heineke, 1998). Therefore, as stated by marketing scholars, research regarding the design of SDS should consider the link between marketing and operation management

theories, i.e. the information needs of managers related to waiting lines, and customer satisfaction with waiting times (Hill et al., 2002; Evangelist et al., 2002).

The aim of the paper is to study some typical configuration of a SDS with waiting lines (i.e. post-offices, banks, emergency rooms, etc.), where the customer, before experiencing the service, is in queue (e.g. physical queue in a store or virtual queue, on the phone and so on). In this kind of system, the customer satisfaction is mainly influenced by the waiting experience, i.e. by the perceptions concerning how much time is spent waiting for the services (i.e. pre-process wait, in-process wait, etc.), and what kind of waiting experiences the customer received (i.e. solo waits, distractions, information about the expected time to be served, etc.) with respect the value provided with the services. In order to study overall performance of such a system advanced modeling techniques, considering jointly the marketing and operations management theories, are needed. The modeling effort requested by an accurate design of waiting lines can be higher if we have to consider some recent evolutions of these systems: a) the several internet-based interaction channels that can be used to communicate with customer's; b) the technology-enabled delivery channels (i.e. e-services, self-service) that have been added to traditional delivery. For sure, Multi-Channel (MC) and cross-channel SDSs are an emergent issue to be shaped and studied in the perspective of service operations management (Roth and Menor, 2003).

Our paper presents an integrated model to analyze the performance of these kinds of systems using the typical techniques of computer simulation, such as DOE. By using computer simulation, the performance related to different functional logics, such as the rules for prioritizing the customers through the waiting lines and for routing them across the multi-channels, are assessed. This led us to a general knowledge on the system's performance and on the interconnectedness of the operating variables, such as the lengths of queues and the opportunity of delivering cross-channel services with respect to attitudes and behaviors of customers. Customer's behavior and customer-related management policies are modeled supposing customers, when entering into the system, interact with a kiosk and receive their booking ticket. The system performance is assessed with respect to the customer satisfaction, that is modeled as being inversely related to the perceived, not only actual, waiting time (Maister, 1985; Davis and Heineke, 1998).

Therefore, our paper is organized as follows. In the next two sections, the most relevant literature related to the main research fields covered by the study is briefly reviewed: the design aspects of multi-channel SDS and the models that refer customer satisfaction to the customer's waiting experience. Then, we illustrate the conceptual model of the SDS we have taken as a paradigm of this kind of problems. Finally, after presenting the findings from the simulation analysis, we discuss the most relevant results and point out perspectives for future research.

Multi-Channel Service Delivery Systems (MC SDS)

In recent years, a lot of standard front-office processes (i.e. cash withdrawal, hotel booking, etc.) in different service industries (i.e. transportation, banking, postal services, healthcare, etc.) have been partly or totally re-designed to be self-services and to exploit the benefits of automation provided by ATM's and kiosks (Rowley and Slack, 2003). Besides physical facilities, the era of e-service has risen. E-service is defined as "deeds, efforts or performances whose delivery is mediated by information technology (including the Web, information kiosks and mobile devices)" (Rowley, 2006). Moreover, all e-service is predominantly self-service, whether it is delivered through a web page on either a PC, or a mobile device, or a kiosk. The combination of e-services and standard service processes usually led to the implementation of a MC SDS. In a MC SDS, different channels (i.e. physical facilities, toll-free numbers, internet live-chats, e-mail address, etc.) can be combined to complete the delivery process. The information flows may require or not co-location and/or synchronicity among the customer's and the service provider's interaction processes. As stated by Sousa and Amorim (2010), the delivery channels "provide visible interfaces to customers who use them to engage in service processes". MC SDS has been addressed in literature from different points of view. A first research stream deals with the frameworks that guide the development of the channel interfaces. With respect to these issues, the Technology Acceptance Model (TAM) (Davis, 1986) identifies the relationship between intention-to-use and attitude toward use of a system, in order to predict the system use on the basis of its ease-of-use and its usefulness (i.e. perceived value). TAM was used as the basis for modeling barriers and enablers to the adoption of Self-Service Technology (SST). A second research stream aims to understand the decision process that leads the consumer to choose a specific channel to receive her/his services. It is then clear that the implementation of a MC SDS could be simplified in terms of design, modeling the management rules, etc., if the designers could predict what drives consumers to a specific channel, and which channels are preferred by which type of customers. On this topic, Dabholkar and Bagozzi (2002) verify both the moderating effect of costumer's traits, such as self-efficacy and need for interaction, and of some situational factors, such as the perceived waiting time and the customer's anxiety, on the decision to choose a specific channel. Other research streams are focused to finding the rationale for customer dis/satisfaction in the SST encounter. Meuter et al. (2000) state that customer is satisfied about SST when it solves an intensified need, saves time and delivers performance. Instead, customer is very intolerant of technology/process failures. As Sousa and Amorim (2010) state, "most of the research to date has been carried out from a marketing perspective, focusing on customer-related factors, i.e., factors which influence the design of the front-office processes with the goal of meeting customer requirements […]". Then, these authors solicit to design MC SDS in an integrated way: under an interlinked marketing-operations management perspective, customers' flows between channels (kiosks or personnel) should be balanced in order to maximize the customers satisfaction, to satisfy operational constraints, considering customer needs and attitudes, to minimize the system costs.

Waiting experience

Customer satisfaction depends not only on the level of objective performance delivered by the service provider, but also on the customer-perceived value of the service. Beyond the objective level, is therefore necessary to understand why customers value a particular set of offerings (Roth and Menor, 2003). Satisfaction depends on perceptions (perception approach) or on the difference between perceptions and expectations (disconfirmation approach) (Davis and Heineke, 1998). At the end of a waiting experience for receiving a service, the customer, taking into account the opinion formed on the basis of either past experiences and/or word of mouth, evaluates the overall service quality in terms of the encounter with the employees (courtesy, empathy, etc.), as well her/his perception about the time spent in waiting (Houston et al., 1998). According to several authors, customer satisfaction is inversely related to the perceived waiting time: the longer customers wait, the less satisfied they are (Davis and Heineke, 1994). Moreover, among the different moments of waiting, pre-process in-process or post-process wait, is the first to be more unpleasant and therefore more closely related to the dissatisfaction (Maister, 1985; Taylor, 1994). A long wait in queue, in addition to increasing customer dissatisfaction of served customers, lead others customers to balk or renege. As stated by Pazgal and Radas (2008), it is more important for a firm to reduce the reneging rather than the balking effect: a customer who decides not to join the line (*balking*) does not waste her/his time and probably will return later; a customer, instead, who joins a line but leaves without being served (*reneging*) wastes her/his time and makes the line seem longer for those contemplating joining; as a matter of fact, reneging customers may prompt some others to balk. Several marketing researches were carried out to understand which factors influence the perception of a customer about the time spent in a queue. Starting from the seminal works written by Hornik (1984) and Maister (1985) the relevant factors have been classified as pertaining to:

- the firm's or the customer's degree of control (Davis and Heineke, 1994);
- their origins, distinguishing in individual or situational (Durrande-Moreau, 1999);
- when they acts (i.e. before or during the wait) (Durrande-Moreau, 1999).

Not controllable factors are not so interesting by an operations management point of view. Some example of not controllable individual factors are the service value for the customer, the quality expectations, the mood, the hurry and the ability to fill the waiting time (e.g. talking, reading or listening music). On the contrary, controllable factors can be used to moderate the perceptions about the service provider's performance. Examples are the perceived fairness of the wait, the length of queue and the distractions or information that are provided by the service provider ("manipulated environment"). Regarding the perceived fairness, in its seminal work Maister (1985) states that the feeling that somebody has "cut in front" of you, causes even the most patient customer to become furious. Being the first-come-first-served the only socially accepted rule, any perception of unfairness in queue management leads to an increase in the perceived waiting time, as well as an increase in the number of customers ahead (Heuts, 2009). Even the effect of providing, in a manipulated way, distraction and information about the waiting, has been studied by various researchers. With respect to the manipulation of the service environment, as the introduction of TV, music or readings in the waiting room, results seem to depend on the type of wait. Taylor (1984) states that distracters help reducing the perceived waiting time during a delay in an airport, while Pruyn and Smidts (1998) states that TV does not modify the perception of wait in a hospital waiting room. Information about the length of the queue (when not visible) and/or the waiting time do not affect for short waiting (i.e. less than 5 minutes) while contributing to have a clear perception of the objective waiting time for longer waiting (i.e. higher than 15 minutes) (Hui and Tse, 1996). If announcements help the customers when the situation is critical, they could even have some negative effects because those who usually under-estimate their waiting time, are provided with a greater awareness of the time spent.

Conceptual modeling of MC SDS

According to Bertrand and Fransoo (2002), this work pertains to axiomatic-quantitative research, whose purpose is to explain the behaviour of real life operational processes through quantitative modeling. Namely, we addressed different policies for routing the customers waiting for services, taking into account theories from both operations management and marketing literature, in order to develop the conceptual model of the system that have been used for simulation analysis. The use of computer discrete-event simulation is justified by the complexity of the problem under assessment (Church and Newman, 2000). First and second order dynamics (e.g. routing policies, thresholds, prioritizing customers based on their profiles) are so complicated to not allow analytical formulation of the relationship among variables.

The conceptualization and model development has been performed on the basis of data and information made available by a major Italian company of postal services. However, the modeling of front-office MC SDS is quite general and is therefore well suited to other contexts such as, e.g. banking, financial or insurance services.

We consider a typical 2-tier structure front-office SDS that provides transactional, rather simple services. The 1-tier is represented by the entry points of customers, i.e. touch screen interactive kiosks (named "totem") through which customers compose their service cart and get their booking ticket. Ticketing is used to rule the customer's position in the 2-tier queues, according to the customers' status/profile, to their priorities and their entry times. Customers can authenticate themselves through the totem, in order to act as "profiled customers". In this case, they will be assigned a higher priority. The 2-tier is represented by the counter personnel and by automated kiosks that deliver the service mix.

Three different types of services, namely A, B, C, have been considered to be freely composed in a service cart. A is the service most requested, either alone or in combination with B and/or C. As a result, any possible combination of services (i.e. A, B, C, AB, AC, BC, and ABC) can be requested, even if some service carts are more common than others. The MC SDS has been designed in order to achieve a good performance in either serving the higher volume and most standardized flows and the less frequent mixed combination flows. As depicted in Figure 1, four different queues are used to handle customers waiting for services. Q1 and Q4 are used to line up those who requested service A, while Q2 and Q3 are for service B and C respectively. Each queue is served by a given number of counter staff and kiosks (Q, W, E, and R). While the kiosks (R-type) and the counter staff (Q-type) can deliver only service A, W-type and E-type counters are more flexible and can provide, instead, different services (i.e. B + A, C + A). As a result, R- and Q-type servers may prompt customers solely from Q4 and Q1 respectively, while W- and E-type servers call customers from a main queue (Q2 or Q3 respectively) and a secondary queue (Q1). The customers' flow is regulated by three management rules: i) the routing policy of customers in a queue; ii) the sort order of the queue; iii) the call order at the counter/kiosk. After a customer composed her/his service cart, the system routes her/him in the queue that, at that moment, will let her/him experience the lower waiting time. Expected waiting time is estimated on the basis of the requests already processed for that kind of service. The system routes the customers to the kiosks (Q4) only if they are willing to use SST, and that option is convenient in terms of time savings. If customers require only service B (C), s/he necessarily will be routed to Q2 (Q3). If, instead s/he requests service A, all queues can be selected and the choice will consider the highest time savings. Finally, for composite carts (i.e. A + B, A + C, etc.), the routing takes into account the convenience to wait for a counter staff capable of delivering all the services in the cart, or to split the delivery process into two or more steps. Once the queue is selected, the sort order is regulated by the priority assigned to serve different customers. In facts, each customer that enters into the system, besides the requested service mix, is characterized by: i) her/his status (profiled vs. not-profiled customers); ii) her/his liking to use SST; iii) her/his reneging time (i.e. the maximum time s/he is willing to wait before leaving the queue); iv) her/his balking level (i.e. the maximum number of customers in queue which will discourage her/him to join the queue). To be more precise, the preliminary part of each queue is dynamically reordered with respect to the priority level assigned to customers in line, while the last part (i.e. the three Next-To-Be-Served – NTBS - customers), instead, follows a FCFS rule.

As already pointed out, while Q-type counters (otherwise R-type kiosks) call the next customers from Q1 (otherwise Q4), W-type and E-type counters, instead, call the next customers following a threshold logic: if the number of customers waiting in the primary queue (Q2 or Q3) exceeds the threshold value, they are served with priority, otherwise customers are called from Q2 (Q3) or, alternatively, from Q1, according to the FCFS rule. Customers, who applied for a serving cart that can be provided only by different counters, after they have received the first part of the service, directly enter in the following queues just at the end of the NTBS.

Figure 1 depicts two examples of routing customers. (a) diagram is the case of a customer who requests the service A and does not want to use SST. The totem then routes her/him to Q1. One of counter staff Q, W or E-type, based on the threshold values, will serve her/him. (b) diagram, instead, represents the case of a customer who requests the serving cart B + C. As the services B and C cannot be provided by the same counters, the delivery will be necessarily split into two; thus, customer should wait both for W-type and E-type counters. Based on the expected waiting time, the system decides to route initially the customer in Q3 and, after s/he received service B, s/he is sent to Q2 just at the end of the NTBS.

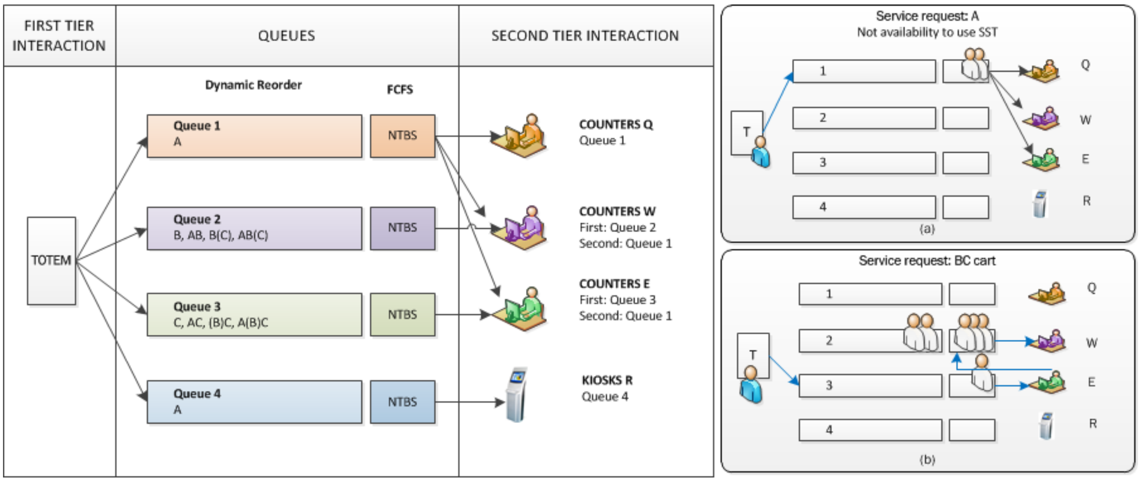

Figure 1 - Conceptual model of the MC SDS (services in brackets are not deliverable by the corresponding counter staff)

Computer simulation and experimental design

The model was coded to be simulated using the discrete event simulation software ARENA®. Real life dynamics are represented through modeling customers' arrivals, routing and waiting experience until the service encounter. We then performed a 2-level Design Of Experiment (DOE), simulating the process of service delivery, to identify which factors are statistically significant to have a major influence on the system performances.

Factors

Five factors, grouped into two categories, were used for this DOE analysis. In the following, we describe the factors and the levels (-, +) they assume in the 2-level DOE. The first category includes factors concerning the queue management policies.

1. <u>Management policy of customer profiled</u>.
   a. Level (-): absence of customer profiling. Each customer is equal and served following the FCFS rule.

b. Level (+): the arrival time (T_{arr}) for profiled customer is decreased by 50% of expected waiting time (EWT) according to equation (1) and then the customer are called following FCFS rule according to time T*.

$$T^* = T_{arr} - \alpha 0{,}5EWT \qquad \alpha = \begin{cases} 1 & if\ the\ customer\ is\ profiled \\ 0 & otherwise \end{cases} \qquad (1)$$

2. <u>Threshold value</u>: number to compare with the number of customers waiting in the main queue for that counter.
   a. Level (-): threshold value is equal to 1. The next customer to be served is called from both the main and secondary queue following the FCFS rule.
   b. Level (+): threshold value is equal to 3 or 8 (respectively for Q2 and Q3). The customers in the main queue are served with priority over those in the secondary one when their number is more than the threshold value.
3. <u>Routing parameter</u>: is the evaluation of the EWT. The lower EWT determines the queue the customer will be routed in, and whether or not the service cart delivery will be split.
   a. Level (-): the number of customers in lines will be the estimate of EWT.
   b. Level (+): knowing the type of service required by each customer, the sum of service times of customers in the different routes divided by the number of active counters, distinguishing between dedicated and shared counters, will be the estimate of EWT.

The second category concerns factors used to characterize individual behaviors and attitudes of customers. Of the customers' attributes reported in the previous section, only two were considered as DOE factors:

4. <u>The percentage of customers willing to use SST;</u>
5. <u>The percentage of profiled customers.</u>

Through the level (-) to level (+), the values increase by 50% in order to vary the population mix.

System performances

Two performances are considered for the system: the average waiting time in each queue (*operational performance*) and the Perceived Waiting Time (PWT) (*service performance*). PWT is linked to the Objective Waiting Time (OWT). The relationship has been considered as moderated by two variables: i) the number of customers in queue (CIQ_i), that represents the number of customers in the queue (as an integer > 0) upon arrival of the i-th customer; ii) the perceived fairness of waiting (PFA_i), as an integer that takes into account the overtaking done (or suffered by) the i-th customer, as a consequence of the other customers routed in that queue with higher or lower priority. In particular:

- $PFA_i = 0$ if the original position in the queue is not manipulated;
- $PFA_i > 0$ if the i-th customer is overtaken;
- $PFA_i < 0$ if the i-th customer overtakes, or if customers ahead of him decide to renege.

We assume that, according to Heuts (2009) and Maister (1985), the higher CIQ_i and PFA_i, the greater PWT respect to OWT. Moreover, the marginal increase in PWT decreases when OWT increases (following the psychophysical literature, see Antonides et al., 2002). For this reason, $OWT_{i,j}$ has been divided into classes *j*. The PWT is then determined according to equation (2):

$$PWT_i = \sum_j PWT_{i,j} \qquad PWT_{i,j} = \left(1 + k_j C_i + h_j F_i\right) \cdot OWT_{i,j} \qquad (2)$$

where:

$$C_i = \begin{cases} 0 & if\ CIQ_i \le 3 \\ \frac{CIQ_i}{\max_i CIQ_i} & if\ CIQ_i > 3 \end{cases};\ F_i = \frac{PFA_i}{\max_i PFA_i};\ (k_j, h_j) = \begin{cases} \left(\frac{1}{3}; \frac{2}{3}\right) se\ j = 1 \\ \left(\frac{1}{6}; \frac{1}{3}\right) se\ j = 2 \\ \left(\frac{1}{10}; \frac{1}{5}\right) se\ j = 3 \end{cases}$$

Findings from the simulation analysis

The DOE was performed for a total of 32 combinations. In order to have statistics built on samples as unbiased estimators, each combination was replicated 10 times. All statistical analysis were carried out using MINITAB®. The first assessment was aimed to identify the factors that have a statistically significant influence on the response variables. The most valuable findings can be described as follows.

OWT depends mainly on both threshold values (T) and the percentage of customers willing to use SST (K). One exception is in the waiting statistics of the kiosk queue, that presents, as a main factor, the routing logic adopted (R), besides the more obvious K-factor. The influence of T and K is easy to explain: the threshold, determining the choice of calling the customer waiting in a queue rather than in another, inevitably impacts on the waiting times of different queues; the same way, if willingness to use kiosks increases, the e-service channel will be more exploited, thus leading to different waiting times in the other channels. R-factor affects only the kiosks and the overall average waiting time in queue; in other terms, the effect of a more precise criterion of routing customers is not visible when looking at a single queue rather than at the system level. In the case of kiosk, the relevance of the R-effect on waiting time is explained by the fact that customers that intend to use the kiosk will be routed to it if, and only if, that is the fastest option.

The customers’ perception about waiting time is analyzed distinguishing between profiled and not profiled customers. We considered the percentage of customers that have a PWT equal or less than OWT, and those with PWT greater than OWT. The main factors that have been considered, are: i) the management policy of customer profiled (P); ii) the threshold values; iii) the percentage of customers willing to use SST. Considering the profiled customers, as expected, the main factor is P: in fact, through level (-) to (+) the sort order of queue enables those categories of customers to proceed more quickly through the queues. As a result, CIQ is reduced and PFA increased. The other two factors, instead, are significant because they affect the OWT for the same reasons stated above. The perception of not profiled customers instead, is mainly influenced by OWT and then by T and K factors. In this case, P appears as a combined effect with K. We note that the effect of the factor related to the percentage of profiled customers is never statistically significant: it was therefore neglected in further analysis.

From the outcomes of the analysis of the main factors, it is possible to determine which combination of factors’ levels yields to less waiting time in line. If the percentage of people willing to use SST increases, the waiting time in all queues at counters is reduced. The waiting time for the kiosk, instead, due to the increased demand, will be longer. All customers who require, at least, service A will use, in fact, more likely, ATMs. The R-factor at level (+), instead, reduces the waiting time for the kiosk because, improving the estimate of EWT, customers will be routed to it only if that choice is really convenient. The introduction of threshold values (level (+)) results in increased waiting time in Q2 (or Q3), because these queues will be served only when the threshold value is exceeded. The management policy of customers profiled does not affect the waiting time even if at level (+) contributes to its reduction, interacting with the K-factor. As a matter of fact, this analysis shows that the goal to reduce the customers' OWT can be achieved by setting all factors at level (+), except for the threshold value. This configuration disadvantages only Q1. However, Q1 is used for customers waiting for service A, that is also provided by kiosks. Investing resources to encourage the use of these devices, it is possible to offset this negative effect and reduce OWT as well. Anyway, it is interesting to assess the impact these decisions have on PWT. Regarding the profiled customers, the main effect is given by the management policy of the profiled customers, and by the interactions between this factor with the threshold value and with the percentage of customer willing to use SST. Setting the factors' levels as above, the PWT is equal or less than OWT, so customers appear to be satisfied. In this case, not profiled customers, for which PWT was mainly influenced by factors T and K, have a minimum level of dissatisfaction as well (PWT > OWT); in fact, as they cannot take advantage of profiling, they will be the more satisfied the lower is OWT. All these considerations vary depending on the percentage of customers willing to use SST, but remain unchanged compared to the percentage of customers profiled.

Concluding remarks

Some considerations emerge from this paper. This work is axiomatic and therefore qualitative: it is not in the authors' scope to present a clear set of innovative management policies. Instead, we emphasize the importance of establishing an integrated approach to assess the different scenarios, not only to take into account the management logics, but also how people interact with these systems. Our modeling is a first attempt to merge these aspects. One of the major features of this work is the customers profiling technique. The need for companies to know their customers is well known. It is therefore interesting to investigate the impact of profiling and how the lack of action on the overall renovation of the management logics leads to reduce the value it brings. On the other hand, with these models we can try an economic optimization of the provider's operations. The provider, in fact, establishes trade negotiations with customers, selling the time savings for those who are profiled and maybe rewarding for the undue delay for those who are not profiled. Regarding the operational issues, the importance of properly ruling the queues is confirmed. On one side, routing customers to the right queues must be determined as accurately as possible, to balance the overall waiting time. On the other side, the call order based on thresholds should be considered as an effective management policy to have the workload leveled in an easy way. However, the introduction of ATMs can reduce largely these benefits: if people become more and more acquainted to use kiosks, then implementing complex balancing rules can be counterproductive.

With respect to future research, the model can be upgraded in several ways. For example, the impact of dynamic counters' shift scheduling can be explored. Another issue to be investigated is related to the use of experimental data (i.e. achieved from field surveys) as an input to the simulation.

References

Antonides, G., Verhoef, P.C., van Aalst, M. (2002), "Consumer perception and evaluation of waiting time: a field experiment", *Journal of Consumer Psychology*, Vol.12, No.3, pp.193-202.

Bertrand, J.W.M., Fransoo, J.C. (2002), "Operations management research methodologies using quantitative modelling", *International Journal of Operations & Production Management*, Vol.22, No.2, pp.241-264.

Carlzon, J. (1989), *Moments-of-Truth. New Strategies for Today's Customer Driven Economy*, Harper & Row, New York, NY.

Church, I., Newman, A.J. (2000), "Using simulations in the optimisation of fast food service delivery", *British Food Journal*, Vol.102, No.5/6, pp.398-405.

Dabholkar, P.A., Bagozzi, R.P. (2002), "An attitudinal model of technology-based self-service: moderating effects of consumer traits and situational factors", *Journal of Academy of Marketing Science*, Vol.30, No.3, pp.184-201.

Davis, F.D. (1986), "Technology acceptance model for empirically testing new end user information systems: theory and results", *Doctoral dissertation*, Sloan School of Management, Massachusetts Institute of Technology.

Davis, M.M. Heineke, J. (1994), "Understanding the role of the customer and the operation for better queue management", *International Journal of Operations & Production Management*, Vol. 14 No. 5, pp.21 - 34.

Davis, M.M. Heineke, J. (1998), "How disconfirmation, perception and actual waiting times impact customer satisfaction", *International Journal of Service Industry Management*, Vol.9 No. 1, pp.64-73.

Durrande-Moreau, A. (1999), "Waiting for service: ten years of empirical research", *International Journal of Service Industry Management*, Vol.10 No.2, pp171-189.

Evangelist, S., Godwin, B., Johnson, J., Conzola, V., Kizer, R., Young-Helou, S., Metters, R. (2002), "Linking marketing and operations: an application at Blockbuster, Inc.", *Journal of Service Research*, Vol.5, No.2, pp.91-100.

Heuts, N. (2009), "The moderating role of customer and situational characteristics", Master thesis Maastricht University School of Business and Economics.

Hill, A.V., Collier, D.A., Froehle, C.M., Goodale, J.C., Metters, R.D., Verma, R. (2002), "Research opportunities in service process design", *Journal of Operations Management*, Vol. 20, pp.189 - 202.

Hornik, J. (1984), "Subjective vs. Objective time measures: a note on the perception of time in consumer behaviour", *The Journal of Consumer Research*, Vol.11, No.1, pp.615-618.

Houston, M.B., Bettencourt, L.A., Wenger, S. (1998), "The relationship between waiting in a service queue and evaluation of service quality: a field theory perspective", *Psychology & Marketing*, Vol.15, No.8, pp.735-753.

Hui, M.K., Tse, D.K. (1996), "What to tell consumers in waits of different lengths: an integrative model of service evaluation", *The Journal of Marketing*, Vol.60 No.2, pp.81 – 90.

Maister, D. (1985), "The Psychology of Waiting Lines" in John Czepiel, Micheal Solomon, and Carol Suprenant, Eds., *The service encounter*, Lexington, MA: Lexinghton Books, 113-123.

Metters, R., Marucheck, A. (2007), "Service Management – Academic issue and scholarly reflections from operations management researchers", *Decision Sciences*, Vol. 38, No.2, pp.195-214.

Meuter, M.L., Ostrom, A.L., Roundtree, R.I., Bitner, M.J. (2000), "Self-service Technologies: understanding customer satisfaction with technology-based service encounters", *The Journal of Marketing*, Vol.64, No.3, pp.50-64.

Parasuraman, A., Zeithalm, V.A., Berry, L.L. (1994), "Reassessment of expectations as a comparison standard on measuring service quality: implications for further research", *Journal of Marketing*, Vol.58, No.1, pp.111-124.

Pazgal, A.I., Radas, S. (2008), "Comparison of customer balking and reneging behavior to queueing theory predictions: an experimental study", *Computers & Operations Research*, Vol.35, pp.2537-2548.

Pruyn, A., Smidts, A. (1998), "Effects of waiting on the satisfaction with the service: beyond objective time measures", *International Journal of Research in Marketing*, Vol.15, pp.321 – 334.

Roth, A. V., Menor, L. J. (2003), "Insights into service operations management: A research agenda", *Production and Operations Management*, Vol.12, No.2, pp.145-163.

Rowley, J. (2006), "An analysis of the e-service literature: towards a research agenda", *Internet Research*, Vol.16, No.3, pp.339-359.

Rowley, J., Slack, F. (2003), "Kiosk in retailing: the quiet revolution", *International Journal of Retail & Distribution Management*, Vol.31, No.6, pp.329-339.

Sampson, S.E. (2000), "Customer-supplier duality and bidirectional supply chains in service organizations", *International Journal of Service Industry Management*, Vol. 11 No.4, pp.348-364.

Sousa, R., Amorim, M. (2010), "Operational criteria for the design of front-office processes in multi-channel service delivery systems", *Proceedings of the 17th European Operations Management Association International Conference*, Porto, Portugal.

Taylor, S. (1994), "Waiting for service: the relationship between delays and evaluation of service", *Journal of Marketing*, Vol. 58, No. 2, pp.56-69.